Volume 56, Part 1 February 1965

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY

ARTHUR SUMMERFIELD

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W.1

AMERICAN BRANCH:

32 EAST 57TH STREET, NEW YORK, N.Y. 10022

25s. net $4.25

Preparation of Papers for the British Journal of Psychology

In submitting papers for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.
It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references, and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. 56 (1965) onwards.
A brief summary, suitable for use as an abstract, should precede the article.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.
Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.
The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.
Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–48.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *L'Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.
The Oxford Dictionary is taken as the authority on spelling.
For other general guidance, authors are referred to *The Journal of Physiology: Suggestions to Authors*. Cambridge University Press. Price 2*s*.

// PSYCHOLOGICAL REVIEW

Edited by RICHARD L. SOLOMON, *University of Pennsylvania*

Vol. 72, No. 1 **January 1965**

CONTENTS

PUBLISHED BIMONTHLY BY THE

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1200 SEVENTEENTH STREET, N.W.
WASHINGTON, D.C. 20036

Subscription price: $10.00 (foreign $10.50) $2.00 per issue

JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR

A NEW SERVICE FOR JEAB SUBSCRIBERS

As the research literature on behavioral processes continues to accumulate at an ever-increasing pace, the task of locating previous material on a given topic becomes correspondingly difficult, frustrating and time consuming. As a step toward meeting this problem, the *Journal of Experimental Analysis of Behavior* proudly announces a new service that has been enthusiastically received by its subscribers. In addition to its volume index, the Journal is now distributing standard 3 in. × 5 in. reference cards for all articles and notes published in its pages. These cards provide suggested headings for appropriate classification of each item, with space for the reader to add his own. The cards also include the author, title, volume, pages, and an abstract of the contents. Subscribers to the Journal receive these cards free of charge.

Please address subscription orders to:

KAY DINSMOOR, JEAB
DEPARTMENT OF PSYCHOLOGY, INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405

A price list of back issues is available on request

Brit. J. Psychol. (1965), **56**, 1, *pp.* 1–13
Printed in Great Britain

BINOCULAR BRIGHTNESS AVERAGING AND CONTOUR INFORMATION

By W. J. M. LEVELT
Institute for Perception RVO–TNO, Soesterberg, The Netherlands

Binocular brightness averaging has been investigated under two conditions: with identical contour information in both eyes and with different contour information.

Equibrightness curves are presented for the simple case, in which right and left test fields are identical in pattern but different in luminance. These curves are for the most part linear; i.e. if the weighted sum of left and right luminance is constant, the same binocular brightness impression is produced. The sum of the weighting co-efficients is unity (law of complementary shares). In the absence of eye dominance, the weights are equal; otherwise a correction for eye dominance must be made.

If monocular contour information is present in one test field, brightness averaging remains linear, but the weight for that eye increases at the cost of the weight for the other eye. In a region close to a monocular contour (within 1° of visual angle), the weight approaches unity, so that binocular brightness in this region is dependent upon the luminance in one eye only.

A suggested explanation of Fechner's paradox is given, and the implications of the approach for the mechanism of binocular rivalry are considered.

The present paper describes a number of experiments on binocular brightness averaging, and on the influence of contour information in one eye upon binocular brightness interaction.

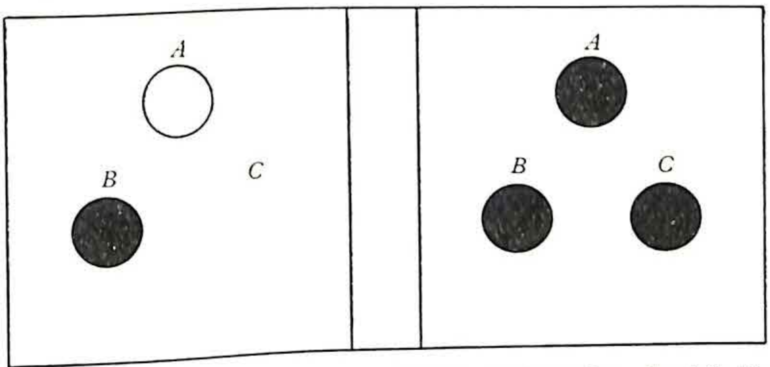


Fig. 1. Stereoscopically the disk *A* seems considerably brighter than *C*, while *B* and *C* do not appear very different in brightness.

Fig. 1 illustrates binocular brightness interaction. If one looks stereoscopically at the three disks *A*, *B* and *C* in Fig. 1, and compares the brightness of these disks, most people have the impression that *A* is considerably brighter than *C*. For the centre of the disks, however, the stimulation of the eyes is identical for *A* and for *C*; the brightness impressions are nevertheless different because of the contour in the left field of *A*. On the other hand, most observers do not see a difference in brightness between *B* and *C*, notwithstanding the fact that the stimulus-object on the left is darker for *B* than for *C*. It thus appears that different stimulation of the eyes may result in nearly equal impressions of brightness, when there are differences in contours.

The literature on binocular brightness interaction is of two main kinds. In studies of one kind, brightness interaction has been investigated as related to absolute

thresholds of vision; the question then is whether the absolute threshold in one eye is dependent upon stimulation of the other eye on a corresponding area. Generally, no interaction of this sort has been reported, or only a very slight one (Graham, 1930, 1931; Crawford, 1940; Pirenne, 1943; Galifret, 1954). Bouman (1955), in his summary of the evidence, arrived at the same conclusion, but showed that there is a large rise of threshold for one eye within a short time (about 0·4 sec.) after the onset of stimulation of the other eye.

Studies of the second type have been concerned with brightness averaging, i.e. the apparent binocular brightness when the left and the right eye are exposed to different luminances (Panum, 1858; Fechner, 1861; Aubert, 1865; Hering, 1865; Sherrington, 1908; De Silva & Bartley, 1930; Graham, 1931; Fry & Bartley, 1933; Ivanoff, 1947; Fiorentini & Radici, 1961; Treisman, 1962). Two phenomena of brightness interaction which at first sight are contradictory emerge from these studies. First, there is 'Fechner's paradox': an object of some luminance is observed binocularly but with a neutral filter in front of one of the eyes. If this eye is then closed, the brightness of the object seems to increase, although the total amount of stimulation is decreased. The second phenomenon is called brightness summation: unequal stimulation of the two eyes produces a binocular brightness which is intermediate between the monocular impressions. A better description of this phenomenon seems to be brightness *averaging*. The term summation has come into use, because the increase of luminance in one eye raises the binocular brightness impression.

Brightness averaging has been studied rather extensively in the past century by Panum (1858), Fechner (1861), Aubert (1865), Hering (1865) and others. Sherrington (1908) and De Silva & Bartley (1930) have reported systematic measurements during this century. Other studies have been confined to demonstrations (Fry & Bartley, 1933; Ivanoff, 1947). The most extensive measurements are those by Aubert (1865, p. 286). However, his procedure may be criticized since he compared binocular stimuli with a monocular one, the luminance of which had to be adjusted by the observer by means of an episcotister. Given the Fechner paradox, it would be better if a binocular test stimulus were compared with a binocular comparison stimulus. This is what Sherrington did (1908, p. 375), but he measured the binocular brightness of five luminance pairs only. Moreover, Sherrington had the test field for one eye constant while the luminance for the other eye varied, and the observer had to adjust the comparison field with equal luminance for both eyes. As a consequence the resulting data do not give an equibrightness curve, i.e. different pairs of left and right luminances which are perceived as equal to a comparison field with constant and identical brightness for both eyes. The same applies to De Silva and Bartley's study. They kept the right field at a constant value, and used seven luminance values for the left test field. Their comparison field usually had a large spatial separation from the test field which was presented simultaneously. This introduces extra variability into the matching behaviour. In the present investigation, a number of equibrightness curves have been determined with a new matching technique.

Experiment I

Binocular brightness averaging

A sensible way of collecting data for equibrightness curves is to have the observer adjust a binocular test field in which the luminance for one eye is fixed by the experimenter, and that for the other eye adjusted by the observer until it produces the same brightness impression as a binocular comparison field with equal luminances for both eyes. A further requirement is that comparison and test fields are projected on corresponding retinal areas, which is the case when both stimuli are centrally fixated.

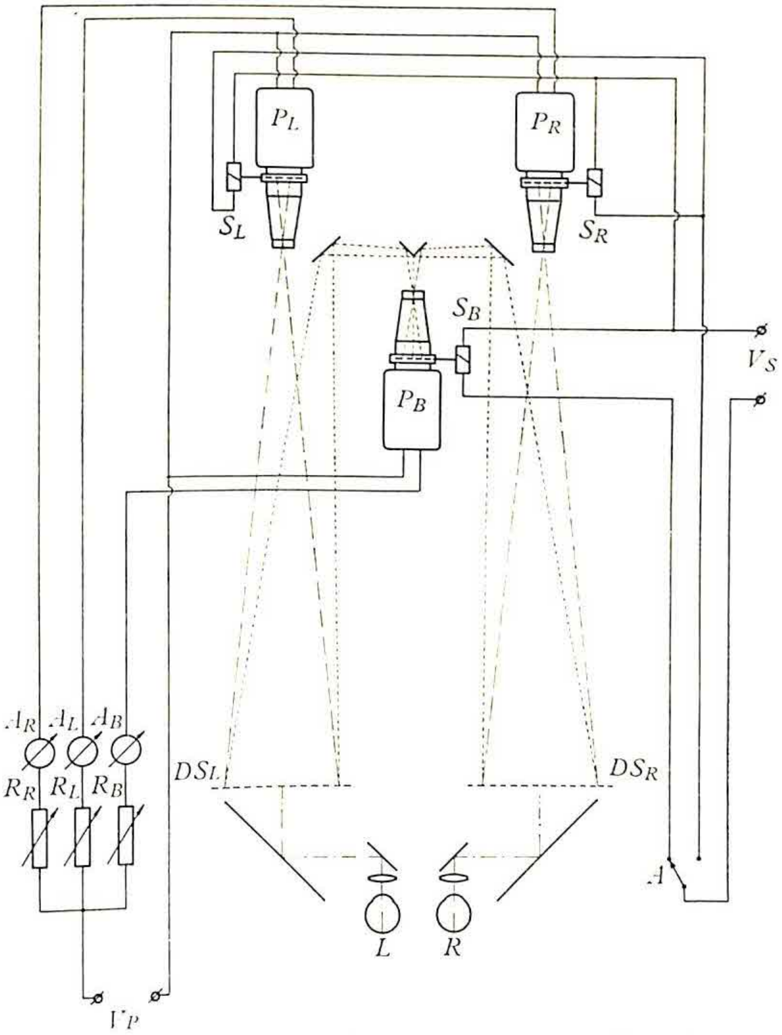


Fig. 2. Arrangement of the apparatus (see text, Expt. I).

Method

Apparatus

The apparatus is schematically represented in Fig. 2. The light sources for the test field are two 150 W. Prado projectors P_L and P_R, the current supply of which can be regulated by two variacs R_L and R_R. They throw light on the diffusing screens DS_L and DS_R in front of which masks can be mounted with holes of the proper size. The projector P_B is used for the comparison field where equal luminance for both screens is needed. This is achieved by splitting the light beam from P_B by means of a prism and two surface mirrors. Alternation of test field and com-

parison field is regulated by a mechanism A, so that if the shutters S_R and S_L are closed, S_B is synchronously opened, and vice versa. The images on DS_L and DS_R are projected to the eyes by means of two surface mirrors and two prisms. Lenses in front of the eyes serve to produce accommodation at infinity. Artificial pupils of 1 mm. diameter are used. The observer's head is supported by a chin rest. Supply voltage was stabilized. Luminances were carefully calibrated, without artificial pupils, to a precision of within 5%.

Procedure

The stimuli which were used in this experiment are shown in Fig. 3. They were circular disks subtending 3° of visual angle against a black background. The luminance of the two disks in the comparison field was set at a chosen value by the experimenter. The luminance of the right test field was increased in little steps, specified below, starting at zero. At every step the observer had to adjust the luminance of the left field until the binocular brightness impression was equal to that of the comparison field. The step-wise increase of luminance in the right test field was continued as long as a match remained possible. Then measurements were made similarly for increasing values of the left test field luminance and the observer had to adjust the right one.

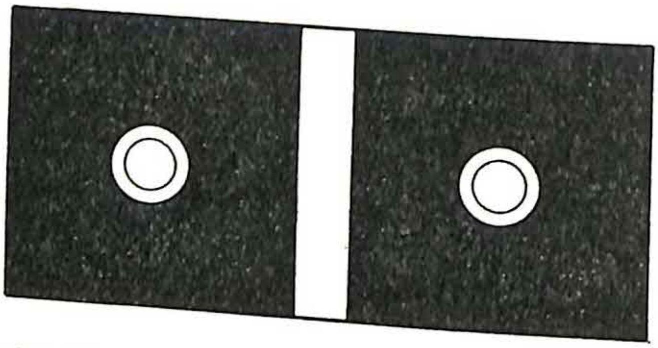


Fig. 3. Stimuli used to determine equibrightness curves. In Expt. I the concentric circles were absent when the data of Fig. 4 were collected, but were present in both monocular fields of both test and comparison fields for Fig. 5; in Expt. II, a circle was present in only one monocular field of both test and comparison fields. The visual angles substended were: circular disks, 3°; concentric circles, 2° in diameter and 3′ in thickness.

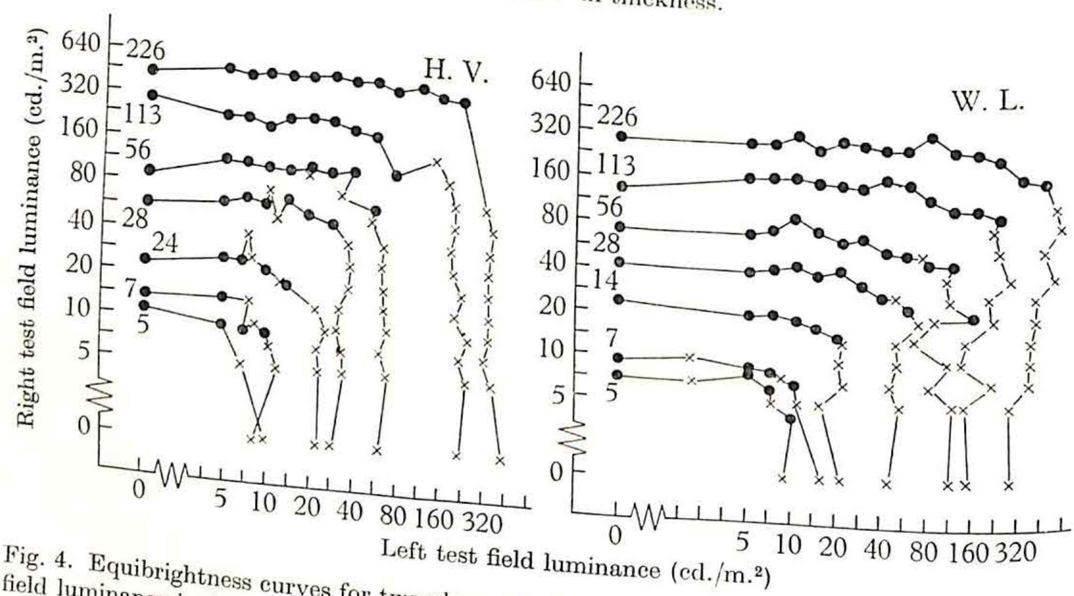


Fig. 4. Equibrightness curves for two observers, H. V. and W. L. The level of the comparison field luminance is shown against each curve. The luminance of the test field was increased in logarithmic steps of 1·4 cd./m.² from 5 cd./m.² upwards; × shows that the observer adjusted the luminance of the left test field and ● that he adjusted that of the right test field.

Series of measurements were obtained for various levels of luminance of the comparison field. The observer could change from the test field to the comparison field and vice versa by manipuating a button, and was free to do so as often as he wanted.

Results

The results, in terms of equibrightness curves, are given in Figs. 4 and 5. Fig. 4 refers to measurements from a series in which the luminance of the field was increased in logarithmic steps of 1·4 from 5 cd./m.^{2} upwards. Fig. 5 gives data for series with linear increases of the test field luminance in steps of 2 cd./m.^{2} from zero. First, as a check on the validity of the matching procedure, one may note that where the observer makes adjustments giving equal luminances for the test fields, their value is in close agreement with that of the comparison field; test and comparison field are identical in these cases.

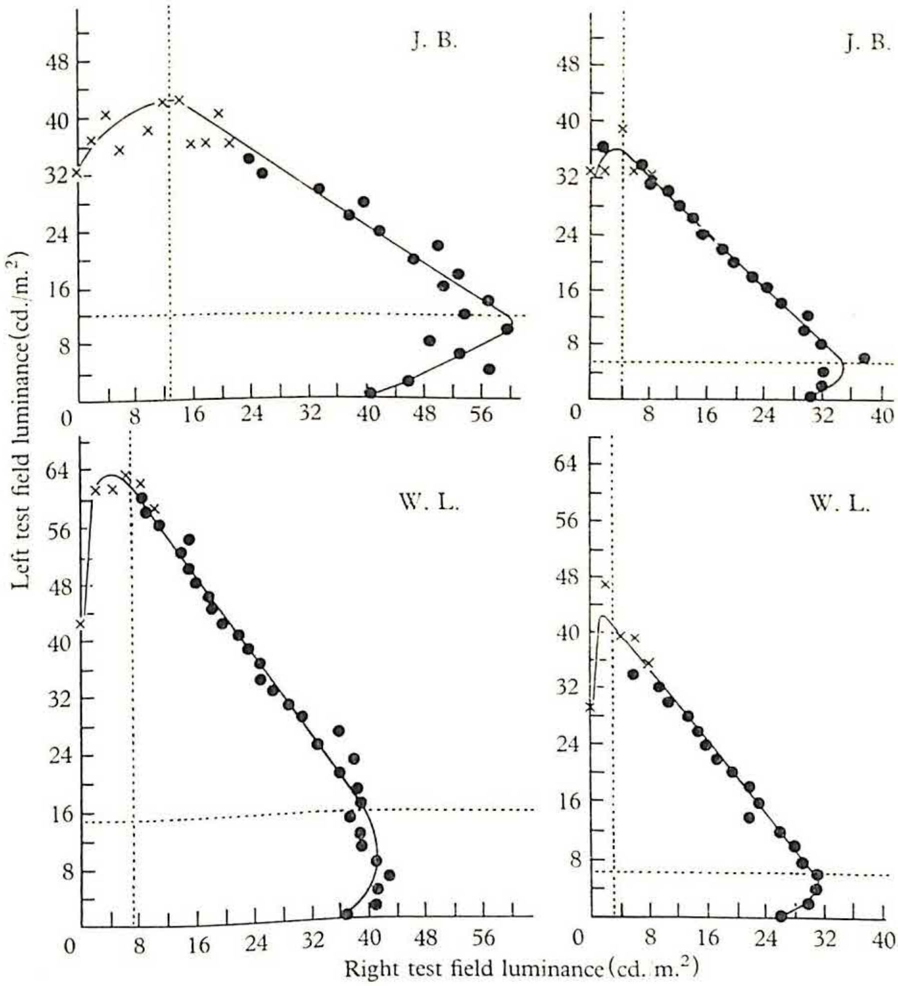


Fig. 5. Equibrightness curves for two observers, J. B. above and W. L. below. Two levels of comparison field luminance were used: 30 cd./m.^{2} on the left in each case, and 20 cd./m.^{2} on the right. The luminance of the test field was increased linearly in steps of 2 cd./m.^{2} from zero; × shows that the observer adjusted the luminance of the left test field and ● that the observer adjusted the luminance of the right test field.

The general trend of the equibrightness curves is clearest in Fig. 5. The functions are linear for test field luminances higher than a value which is indicated in each

figure. The slope of the line is different for the two observers. It is a plausible interpretation that the slope is dependent upon eye dominance; the steep curves for W.L. are for an observer with strong dominance of the right eye; the observer J.B. does not show clear dominance. A difference in sensitivity between the two eyes might be an alternative hypothesis. However, when both subjects were tested on the Haag-Streit Adaptometer during 3 min. after foveal adaptation to 2,000 cd./m.², neither subject showed a difference in foveal sensitivity between the two eyes.

The linear portions of an equibrightness curve can be expressed as: $w_l E_l + w_r E_r = C$. Here E_l and E_r are luminances of left and right test fields, respectively; and w_l and w_r can be interpreted as weighting coefficients which account for eye dominance. For the point $E_l = E_r = E$ we have: $(w_l + w_r)\ E = C$; therefore it is only natural to choose $w_l + w_r = 1$, so that $C = E_b$, the luminance of the comparison field. The fact that the curves are linear, disregarding the tails, implies that binocular brightness averaging can be simply described as an averaging of energies, and thus far there is no special reason to claim that brightness averaging is a matter of averaging of 'sensations'. The latter claim was made by Aubert and later on by Sherrington. Sherrington went as far as to speculate that the sensorium of the right eye is completely separated from that of the left eye. Whatever the truth of this, the present curves suggest that the binocular brightness impression does not result from simple averaging of monocular sensations: it is known from psychophysical studies that monocular and normal binocular subjective brightness are non-linear functions of stimulus energy. Irrespective of whether this is a logarithmic function (Fechner), or a power function (Stevens), or any other non-linear function, the result could never be that binocular brightness is a linear function of its monocular components if sensations were merely averaged. The experimental error in the present measurements evidently tolerated slight deviations from linearity, but even a function of power 1/2 produces a bend, similar in all curves. Therefore, if binocular brightness were a matter of combining sensations, the results suggest that they would have to be combined in a more complicated manner: the resulting binocular brightness would have to be the same as if energies were directly averaged.

It is concluded, therefore, that binocular brightness is constant if a weighted sum of monocular energies is constant; the weighting coefficients add to unity and are constant for an individual observer. This rule is not valid if one of the monocular luminosities is low.

EXPERIMENT II

Monocular contour information

The following experiments were the main experiments of this study and they were undertaken to assess the role of monocular contour information in binocular brightness averaging. The first experiment is an obvious extension of the experiment described earlier. Equibrightness curves were again determined, with the difference that a concentric circle, subtending 2°, was present in one of the monocular fields of both the test field and the comparison field (cf. Fig. 3). The observer was instructed to match the fields for the interior of the circles. The comparison field was kept at 30 cd./m.². In all other respects the procedure was the same as described before.

RESULTS

Fig. 6 shows results for circles in the left fields, and Fig. 7 for circles in the right fields. The curves are again linear, except for the tails, and pass through the point $E_l = E_r = 30$. This means that they may again be described by the function*

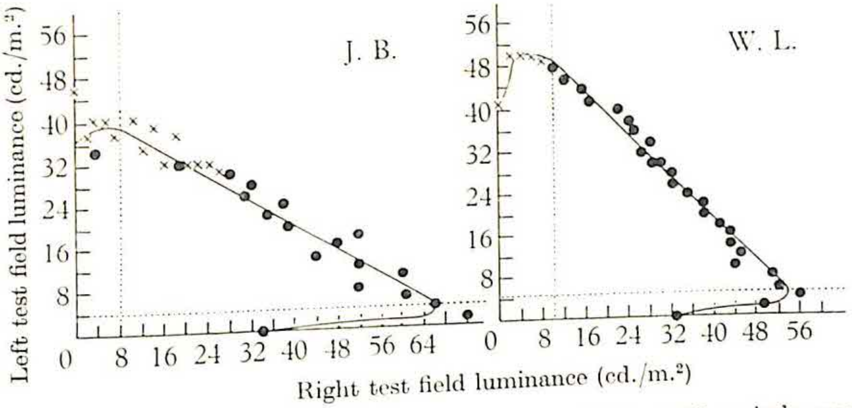


Fig. 6. Equibrightness curves for two observers, J. B. and W. L., with a circle present in the left monocular field of both the test field and the comparison field (see Fig. 3). The luminance of the comparison field was 30 cd./m.².

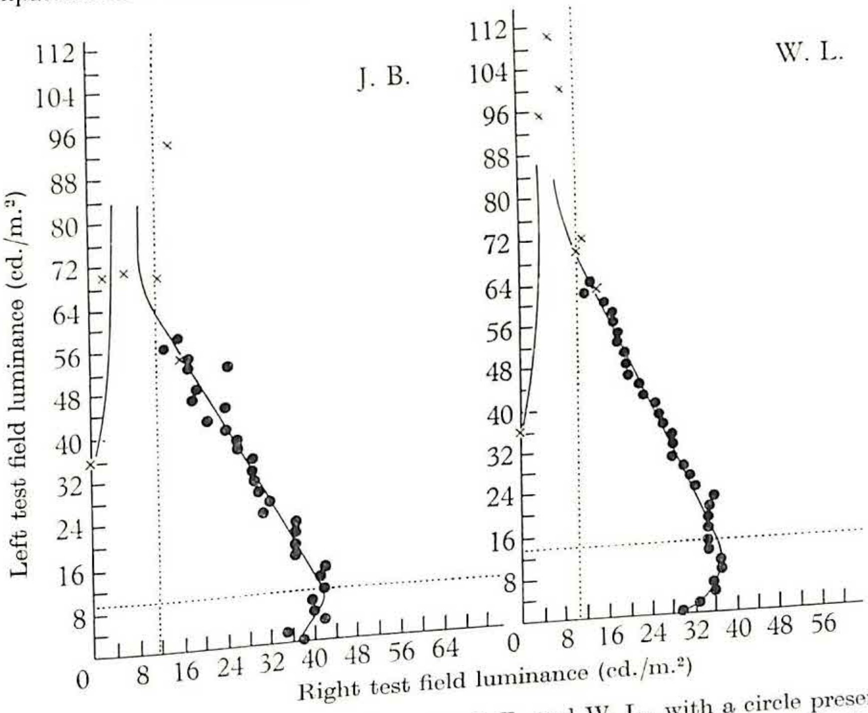


Fig. 7. Equibrightness curves for two observers, J. B. and W. L., with a circle present in the right monocular field of both the test field and the comparison field (see Fig. 3). The luminance of the comparison field was 30 cd./m.².

* The circle in one of the monocular comparison fields serves to make the matching process easier for the observer. The question arises, however, of whether E_b retains the same meaning when C is equated to it. In other words, for equal stimulation of both eyes, does the binocular brightness impression change if a circle is eliminated from one of the monocular fields? This seems very improbable; the assumption was checked by having an observer compare a pair of fields of luminance 30 cd./m.², both containing a 2° circle, to an adjustable pair with a circle in the left field only. Ten adjustments were made; their mean value was 30·6, 0·49. The assumption may be maintained, therefore.

$w_l E_l + w_r E_r = C = E_b$, where $w_l + w_r = 1$. The only difference between the three curves from each observer is in their inclination (see Figs. 5–7). This can be expressed as follows: the weighting coefficient for the eye in which contour information is present is increased relative to the coefficient for the other eye. This may be called the *law of complementary shares*, after Hering.* It simply means that if w_l increases w_r decreases, and vice versa, so that their sum remains unity.

The law enables a comment to be made on the tails of the equibrightness curves, where one of the monocular components has a low value. It is clear that if the luminance of one test field, the left one say, is below threshold, contour information is present in the right field only. Therefore, for these low values of E_l, w_r will increase at the cost of w_l, and so a change in slope of the curves at the tails is predicted (Figs. 4–7). It is not possible to give the precise function of the non-linear parts of the curves; in particular, the luminance values at which the weighting coefficients start changing are difficult to estimate. They are probably not at the threshold value for one eye, but at a value at which the contour information has 'faded out' sufficiently.

EXPERIMENT III

Change in weighting as dependent upon distance from contour

A monocular contour increases the weighting coefficient for the luminance in the corresponding eye. The next question is whether the effect of a contour is local or general, i.e. is there an increase in w for the whole monocular visual field, or is the increase limited to a region in the immediate environment of the contour?

Method

The stimuli for this experiment are shown in Fig. 8. They were four pairs of patterns, subtending 5°, containing circles. These stimuli were used for both the test and the comparison field. The difference was again that in the test field luminances were unequal, 250 cd./m.² for the left eye and 25 cd./m.² for the right eye, whereas in the comparison field they were equal. In this experiment the observer had to adjust the comparison field until the brightness in the centre of the upper circle looked equal for both fields. Four observers served in this experiment. The four different stimulus conditions were given to them in an order which was varied according to a Latin square design.

Results

Table 1 gives the results in terms of the adjusted luminance in the comparison field. An analysis of variance reveals significant differences between the conditions. First, the C pair gave a higher value than the D pair ($P < 0{\cdot}001$). For the C pair the contour was in the more luminous (left) test field, whereas for the D pair it was in the less luminous (right) field; this result therefore confirms the argument that has been advanced that contour information influences the weighting coefficient. Secondly, in both A and B pairs *one* circle is present in the more luminous left field, and one circle in the dimmer right field. If the influence of the contour extends to the whole field, the matchings for the upper circle area in A and B would not be expected to be different. If, however, only the direct environment of the contour is effective, the comparison field for A should be adjusted to a higher luminance than that for B.

* Hering (1865, pp. 308 ff.) suggested this law (*Gesetz des complementären Antheils der Netzhäute am Schraume*) for any kind of binocular interaction, without giving a quantitative specification.

Table 1 shows that the latter is the case ($P < 0{\cdot}005$). The differences between A and C, and B and D do not reach significance. The conclusion is, therefore, that the weighting coefficient is increased only for the immediate neighbourhood of contours, not for the visual field as a whole.

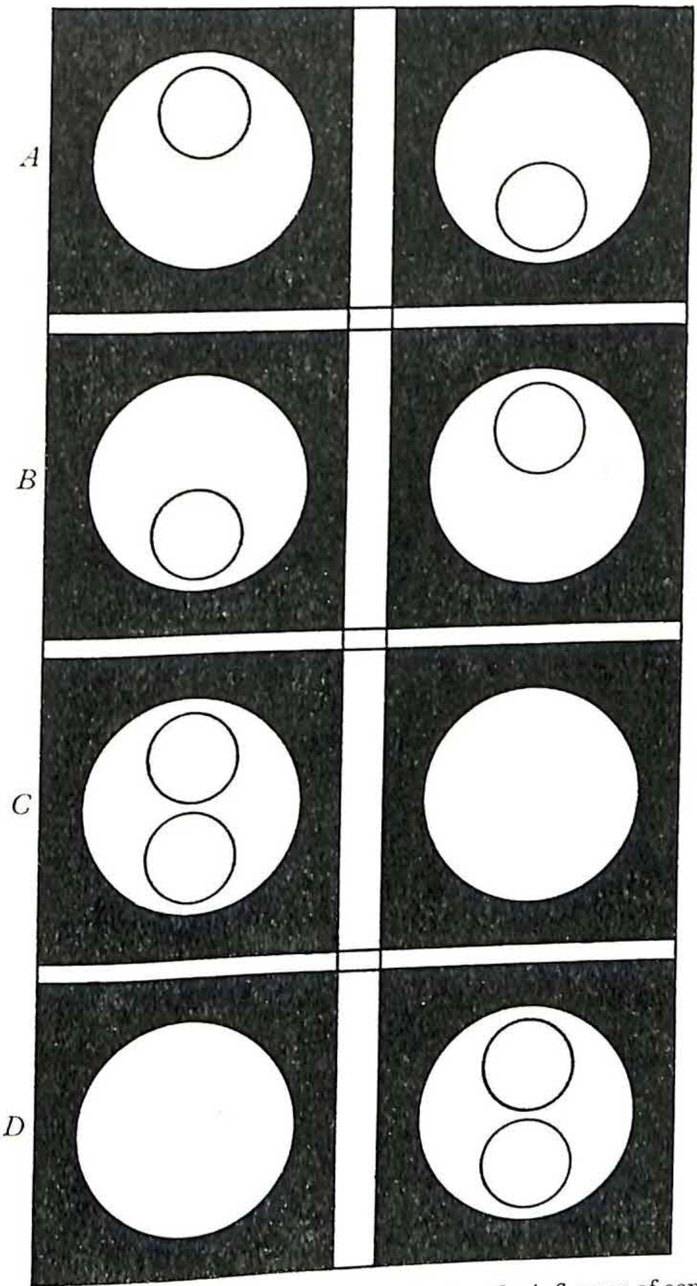


Fig. 8. Four pairs of stimuli which were used to test whether the influence of contours was local or general. The same stimuli were used for both the test and the comparison fields. The diameter of the disks was 5° of visual angle. Test field luminances were always 250 cd./m.^{2} for the left field and 25 cd./m.^{2} for the right field. The observer adjusted the luminance of the area within the upper circle only of the comparison field.

Table 1. *Comparison field adjustments in cd./m.² of four observers for the conditions of Fig. 8, Expt. III*

Observer	Condition A	B	C	D
1	202	53	195	70
2	180	43	190	50
3	180	57	223	80
4	202	117	270	104
Means	191	67·5	219·5	76

EXPERIMENT IV

Spatial amplitude of the weighting variation

The next question follows immediately: to what value does w increase at the fixation point if the distance between the fixation point and the contour is made smaller and smaller? In other words, this question refers to the spatial amplitude of variation in w.

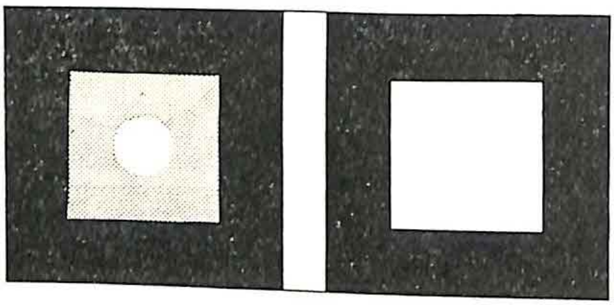


Fig. 9. Stimuli used to determine w_l as a function of field size. The squares were 14° × 14° of visual angle; four sizes of central disk in the left field were used: 1°, 3°, 5° and 7°. The patterns of the left and right test and comparison fields were the same; see text for details of the luminance relations.

Method

The stimulus conditions for this experiment are shown in Fig. 9. The right test field was a square of 14° × 14°. Its luminance was fixed at 100 cd./m.². The right comparison field was identical, but its brightness was adjustable by the observer. The left test field consisted of two parts: a central disk of variable size, with luminance fixed at 12 cd./m.², and a surrounding field of 14° × 14° at luminance 3·7 cd./m.². The left comparison field had the same pattern; the luminance of the central disk was always the same as that adjusted by the observer for the right field, and the luminance of the surrounding area was always 1/3·25 of that of the central disk. The observer had to adjust the comparison field until the brightness of the central disk appeared equal for both test field and comparison field, while fixating the centre of the disk.

Four sizes of disk were used: 7°, 5°, 3° and 1° of visual angle. The conditions were presented according to a Latin square design. Two groups of four observers took part in the experiment.

In this experiment boundaries between two different luminances were used instead of contours on a uniform background. The reason was that it is quite impossible to make a trustworthy brightness match of the area within a contour of 1° on a relatively large, uniformly illuminated background, when instructed to neglect the brightness of this surrounding field. The problems of brightness contrast in this stimulus arrangement were minimized in two ways. The central disk was brighter than the surrounding area, and the brightness contrast effect is only great for a stimulus shown against a more luminous background. Moreover, the ratio between disk and surrounding luminance was equal for test and comparison fields and constant throughout the experiment.

Results

Individual values of w have been calculated for the four disk sizes. The results are given in Table 2. Analysis of variance shows that w_l increases with decreasing diameter of the disks; the regression is significant ($P < 0{\cdot}001$). In Table 2, it is seen that at 1° for observers 3, 6 and 8, w approaches the unit value as closely as adjustment errors permit. In view, moreover, of the increasing tendency apparent in the mean values of w for decreasing size of the disk, the data strongly suggest that in the immediate neighbourhood of a monocularly presented contour, binocular brightness is exclusively determined by the luminance of this monocular field. The strongly localized contour effects in the former experiment (Table 1), compared with the relatively high values of w for all angles in Table 2, suggests that variation of w is limited by the presence of a contour. Furthermore, a boundary may be more effective than a contour.

Table 2. *Values of w_l for eight observers and four sizes of disk*

	Size of disk (visual angle)			
Observer	1°	3°	5°	7°
1	·835	·786	·741	·798
2	·956	·873	·941	·875
3	·997	·981	·974	·968
4	·911	·830	·765	·792
5	·939	·914	·824	·801
6	·990	·900	·881	·864
7	·824	·810	·790	·773
8	1·002	·998	·941	·937
Means	·932	·887	·857	·851

Fechner's paradox can now be understood as follows. If a neutral filter is placed before the right eye, $E_b = w_lE + w_rtE = E(w_l + w_rt)$, where t is the transmission of the filter. If the right eye is closed, w_l equals unity, so that $E_b' = E$. Since $w_l + w_r = 1$, and $t < 1$, it follows that $w_l + w_rt < 1$, and therefore $E_b' > E_b$. Hence the apparent brightness increases if the right eye is closed. This argument implies, however, that the assumption that $w_l + w_r = 1$ remains valid for monocular observation. There is some evidence that this is not true without qualification. In fact, one may interpret the data for $E_r = 0$ (Figs. 6 and 7) as an indication that $w_l + w_r < 1$ there also, since $E_b' < E_l$ (similarly for $E_l = 0$; comparable instances are given by De Silva & Bartley, 1930 as argument for brightness summation). However, two considerations apply.

(*a*) Matching a monocular and a binocular brightness appears to be a rather unstable affair. Day-to-day variability is high. Observers have a feeling that at every new experimental session some arbitrary criterion is chosen, a feeling that is absent for binocular-binocular matchings.

(*b*) To check whether the assumption that $w_l + w_r = 1$ remains valid for monocular observation, a very small test field (or comparison field) has to be used in order to ensure that the field as a whole is sufficiently close to contour information for its weighting coefficient to approach unity. Closeness to contour information may be relevant; in Fechner's original experiments the paradox was absent for filter transmission values closely approaching unity (Fechner, 1861, p. 420), indicating that $E_b' < E$ in the above sense. This follows necessarily if $w_l < 1$, because $E = w_lE_b' < E$.

According to Table 2, w is less than unity at a great distance from the contour, and this was precisely Fechner's situation because he instructed his subjects to look at the blue sky through the filter.

For these reasons it does not seem necessary to reject the law of complementary shares for the case of monocular stimulation. In addition, the explanation of Fechner's paradox would not be invalidated if the law of complementary shares had to be qualified. A good choice of the transmission of the filter might always compensate for minor deviations from the law.

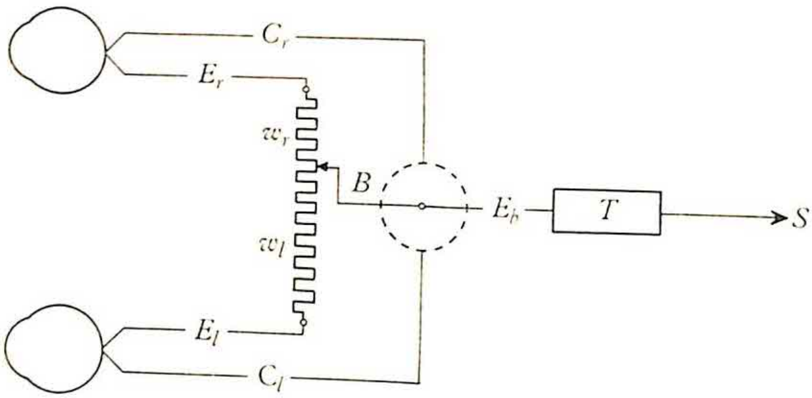


Fig. 10. Schematic representation of the interaction of monocular contours and luminances in binocular brightness averaging; see text for details.

CONCLUSIONS

The general conclusions from this study may be summarized with the help of Fig. 10. This diagram is intended as a mnemonic device, without further pretensions, and is drawn as an electrical flow-diagram. Luminosities E_l and E_r stimulate the eyes and may be thought of as voltages in the electrical metaphor. They are weighted by the balance B, a potentiometer in electrical terms, and their weighted sum E_b is the basis for the binocular impression of brightness; w_l and w_r become variable resistances with a constant sum, representing the law of complementary shares; as a result, the current in the channel past B is $E_b = w_l E_l + w_r E_r$. The weighting coefficients, w, depend upon contour information, as indicated by the channels c_l and c_r to B, which determine the position of the potentiometer. If contour information is similar for both eyes, the weighting coefficients are determined by eye dominance only; in the absence of eye dominance they are both equal to one-half. If contour information is present for one eye only, the weighting coefficient for this eye increases, up to a maximum of unity, with a corresponding decrease in the coefficient for the other eye according to the law of complementary shares. The whole mechanism applies to a small area of the visual field only; for a different area a similar mechanism must be postulated for which the parameters may be different.

If the binocular brightness impression is a function of averaged energies, the 'translation', T, of the weighted energies into a binocular brightness sensation has been simply located in the diagram beyond the averaging process, and T can be imagined as a device with some non-linear amplitude transfer characteristic. At every point

of the visual field, the apparent brightness is thus determined by the left and right eye luminances at that point, and by left and right eye contours near the point.

The implications of the approach for the mechanism of binocular rivalry are considered in conclusion. Rivalry occurs, generally, when non-corresponding contour information is given. Rivalry is not a matter of the monocular visual fields as a whole; parts of one field may enter into rivalry with parts of the other field. In terms of Fig. 10, if non-corresponding contours are present near an area T, a conflict results in the partition of the weights. The contour in the left eye produces a tendency for w_l to increase; the non-corresponding contour in the right eye, in its turn, will produce a tendency for w_r to rise. An increase of both w_l and w_r would evidently violate the law of complementary shares. Apparently this conflict is resolved in such a way that one tendency triumphs over the other for some time, after which the other tendency becomes victorious. The law of complementary shares is thus saved by an alternating process. Evidence for this process will be given in a subsequent report.

The author is grateful to Prof. J. P. van de Geer for stimulating this research and correcting the text, and to Prof. M. A. Bouman for critically reading and discussing the manuscript.

References

Aubert, H. (1865). *Physiologie der Netzhaut.* Breslau: Morgenstern.
Bouman, M. A. (1955). On foveal and peripheral interaction in binocular vision. *Opt. Act.* **1**, 177–83.
Crawford, B. (1940). Ocular interaction in its relation to measurements of brightness threshold. *Proc. Roy. Soc.* B, **128**, 552–61.
De Silva & Bartley, S. H. (1930). Summation and subtraction of brightness in binocular perception. *Brit. J. Psychol.* **20**, 241–50.
Fechner, G. Th. (1861). Über einige Verhältnisse des binocularen Sehens. *Abh. sächs. Ges. (Akad.) Wiss.* **7**, 337–564.
Fiorentini, A. & Radici, T. (1961). Sur un effet d'interaction binoculaire. *Vision Res.* **1**, 244–52.
Fry, G. A. & Bartley, S. H. (1933). The brilliance of an object seen binocularly. *Amer. J. Ophthalmol.* **16**, 687–93.
Galifret, Y. (1954). Contributions à l'étude du problème de l'interaction des stimulations périphériques binoculaires. *C.R. Soc. Biol., Paris*, **148**, 215.
Graham, C. H. (1930). An investigation of binocular summation. I. The fovea. *J. Gen. Psychol.* **3**, 494–510.
Graham, C. H. (1931). An investigation of binocular summation. II. The periphery. *J. Gen. Psychol.* **5**, 311–27.
Hering, E. (1865). *Beiträge zur Physiologie.* Leipzig: Engelmann.
Ivanoff, A. (1947). Au sujet paradoxale de Fechner. *Acad. Sci. Opt. Physiol.* Séance du 19 mai, 1947.
Panum, P. L. (1858). *Physiologische Untersuchungen über das Sehen mit zwei Augen.* Kiel: Schwerst.
Pirenne, M. H. (1943). Binocular and uniocular threshold of vision. *Nature, Lond.*, **152**, 698.
Sherrington, C. S. (1908). *The Integrative Action of the Nervous System.* London: Constable.
Treisman, A. (1962). Binocular rivalry and stereoscopic depth perception. *Quart. J. Exp. Psychol.* **14**, 23–37.

(*Manuscript received 28 January 1964*)

Brit. J. Psychol. (1965), **56**, 1, *pp.* 15–31
Printed in Great Britain

MAZE LEARNING AND VISUAL DISCRIMINATION IN THE WOOD ANT (*FORMICA RUFA*)

BY D. M. VOWLES

Institute of Experimental Psychology, University of Oxford

A technique is described for training ants in a T-maze. The walls of the maze were covered with patterns which served as cues for the correct solution of the maze. These patterns were used to test the visual capabilities of the ant. It was found that ants would discriminate between vertical and horizontal stripes, and vertical and diagonal stripes, but not between diagonal stripes of opposite slope. In an experiment on generalization using vertical stripes of different widths, the ants appeared able to generalize satisfactorily. The implications of these experiments for insect pattern vision are discussed.

The objectives of the work reported in this paper were: (1) to devise a simple method for rapidly training ants to make simple visual discriminations; (2) to use this method for testing current theories of form perception in insects; (3) to investigate the ability of ants to generalize between figures of different sizes in the hope that this might throw some light upon size constancy mechanisms in insects.

Experiments designed to test the ability of animals to distinguish between different stimuli have depended either upon the animals showing a pre-existing preference for a particular stimulus, or upon their ability to learn during the experiment, thus changing their responses in relation to the stimuli used. Experiments on visual perception in bees and butterflies have been concerned with the characteristics of the stimuli provided by flowers, while Wallace (1958) and Ellis and Pearce (1962) have investigated the spontaneous preference of locusts for different shapes of object. It has been clear for some time that one of the most important characteristics of visual patterns to which insects respond is the amount of contour in the figure. It has been suggested that insects distinguish different figures in terms of the quantity of flicker produced by movements of visual contours over the retina of the compound eye. Although flicker is certainly an important characteristic, as Wolf & Wolf (1936) have shown, it is possible (Hertz, 1937) that spatial as well as temporal patterns of excitation might be involved.

Hassenstein (1958) found that in the beetle the perception of movement involves the integration of activities between neighbouring ommatidia. An auto-correlation is performed between the stimulus effects in groups of two or three adjacent ommatidia. Hassenstein has shown that such a mechanism would give good figure ground separation, but has not yet extended his theory to cover shape perception.

In the experiments described below, various visual patterns were used which had different amounts of contour and therefore gave rise to different quantities of flicker as the insect ran past them. Predictions were made as to the relative discriminability of these patterns in terms of the amount of contour they possessed and the flicker they would produce. However, since ants do not fly to flowers it was not possible to use such a training technique as has been successful with bees. Another ability of ants was therefore used, namely their method of homing by means of visual land-

marks. It is well known (for references see Schneirla, 1953; and Vowles, 1955) that ants readily learn to use cues derived from light direction or objects in their vicinity, in order to guide themselves over a particular route to and from their nest. A technique was therefore developed using a simple T-maze through which the wood ant (*Formica rufa*) had to find its way home. It was already believed that this ant is largely dependent on visual cues and this was confirmed in an extensive series of preliminary experiments briefly reported by Vowles (1958). The visual cues used were provided by placing patterns upon the walls of the maze and eliminating other extraneous cues.

It is not known whether insects possess any mechanism for producing size constancy. If they do not possess such a mechanism, then the same figure at different distances should appear to be a different size. If ants are trained to respond to figures of different sizes their ability to learn such a problem should throw some light upon their constancy mechanism. In the final section of the paper such experiments are described.

METHODS

The ants which were used in these experiments were kept in an artificial nest, which consisted of two interconnecting, glass-roofed chambers set in a concrete block. The brood chamber was kept darkened and supplied with nesting material. Food was placed in the outer illuminated chamber. The whole nest was kept fairly humid. The chambers were 3 in. × 2 in. floor space and ½ in. deep. It was found important to have such relatively large chambers with this particular species as groups of ants tend to die, possibly of formic acid poisoning, if kept in too confined a space. The nest was stocked with twenty to thirty workers, brood when this was available, and in the majority of cases a single queen. The condition of the nest is important for the success of learning experiments; if the ants are not thriving, very poor results are obtained.

The T-maze used for training was connected through a glass tube to the outer chamber of the nest, a guillotine-type Perspex door being placed across the tube. Only one ant was allowed through at a time, and ants were prevented from leaving the nest and going out into the maze on their own initiative. The door must also have eliminated at least some olfactory cues from the nest. The glass exit tube sloped gently upwards, since it had been found that ants prefer to walk slightly upwards than to travel level or downwards under these conditions.

The maze itself was constructed of transparent Perspex, and the dimensions are shown in Fig. 1. The alleys were 1 cm. deep. Marked ants were taken individually from the nest and introduced into the maze through a hole in the lid at the entrance to the T. This was done by placing the ant in a specimen tube and waiting for it to climb down into the maze. In the end wall of the T, below and behind the entrance hole, was a swab of cotton wool, moistened with peppermint essence. The peppermint repelled the ants and caused them to move up into the body of the maze. Towards the end of training the peppermint was omitted, and this led to no falling off in performance.

The walls of the maze were lined with white card, to which were fastened striped patterns supplied commercially as 'Zipatone'. The patterns extended all over the walls of the maze, each pattern occupying both walls in one arm of the T, and the adjoining wall in the entrance alley. The two patterns on the long cross wall met at the junction of the arms and the stem.

The maze was screened from its surroundings by a vertical white tubular screen 18 in. high, through which small slits were cut for the exit tube. Usually the specimen tube containing the ant was inverted over the entrance and the tubular screen placed in position immediately afterwards, before the ant had time to descend into the maze. The maze was illuminated by an ordinary bench lamp, with a 60 W pearl bulb, and a round shade, placed level with the top of the cylinder. The centre of the lamp was placed carefully over the junction of the T, to eliminate cues from light direction.

The floor of the maze was lined with white card. Both these linings and the patterned linings on the wall were changed frequently, usually every three days, and wiped daily with a cotton-

wool swab soaked in alcohol, in order to eliminate olfactory cues. The negative alley was closed with a plug of plasticine. It is possible that the difference in smell and appearance between the plasticine and the exit tube could have been used as cues by the ants, but in fact preliminary experiments showed that this was not the case. Three further results support this: first, ants could be set an insoluble visual discrimination, and did not learn to run the maze correctly under such conditions; secondly, after successfully learning a discrimination it was clear from the tracks of the ant that the choice was made in the entrance alley, well before they could see the arms of the T; and thirdly, if the striped linings are replaced by white or other patterns after learning the ants appeared lost, while exchanging them for the similar pattern did not disturb the ants. In fact, odour cues were found to be almost totally unimportant, and even if the precautions to control odour were accidentally relaxed, it was found that the results were unaffected.

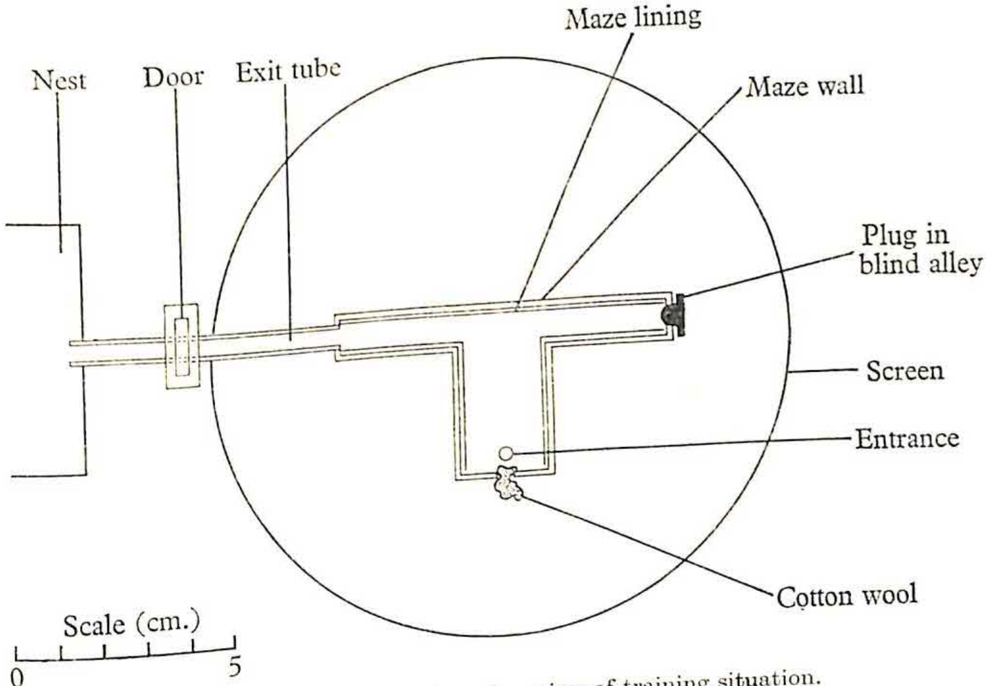


Fig. 1. Diagram showing plan view of training situation.

Ants were marked individually by means of cellulose paint, and eight or ten individuals were labelled in each nest. These were taken individually from the nest at the times of their trials and placed in specimen tubes to await their turn. There was usually a period of between 2 and 10 min. after their removal from the nest and before their entrance into the maze. Ants were run individually, and only one trial per day was given. It was found that two or more daily trials, if regularly given, led to abnormal behaviour on the part of the ant. Similarly, if any punishment (such as an electric shock) was given in the blind alley, the ants also rapidly developed abnormal behaviour such as twitching and circling.

The motivation to leave the maze was presumably to escape from the odour of peppermint and the generally unfamiliar environment, and the reward was merely to return to the nest. However, ants which had returned home often attempted later to emerge again into the maze, but were, of course, prevented from so doing by the closed door: the significance of this is not clear. Ants were observed continuously while in the maze, and their behaviour recorded as described in the section on results.

On successive days the direction in which the ant had to turn to follow the correct pattern was alternated (left, right, left, etc.). It had been found in preliminary experiments that ants rapidly learn to turn in a particular direction relative to their own body-axis, and since in a random sequence a succession of similar turns frequently occurs by chance, this established a turning tendency which conflicted with the visual discrimination and prolonged the learning period considerably. It was felt, therefore, that alternation was better than randomization, and

no evidence has since been obtained that ants can learn to alternate. The experiments on an insoluble discrimination and disorientation after removing the patterned linings showed that indeed they cannot do so at least within the time span of these experiments.

In their early trials in the maze the ants show very considerable back-tracking, and often fail to leave the maze by the exit tube, of which they seemed afraid. Since this initial phase of 'maze-adjustment' (Schneirla, 1941) frequently involves as many as thirty to forty errors per trial, which then lasts 20–30 min., it is tedious and laborious to observe. The experimenter did not require information about this early learning period, and therefore gave all the ants preliminary trials, in a white walled, screened maze, for 3 days before training started. All the experimental ants were placed in the maze at the same time and left with the maze door open for 2–3 hr. Their behaviour was not observed during this period. The direction of turn was alternated on successive days. As a result of this pretraining the ants became familiar with maze conditions and the exit tube. They subsequently showed relatively few errors in the early training situation, and found their way home within 4–5 min.

Learning behaviour and treatment of results

In the early stages of their experience of the maze ants run along very convoluted paths, with many turns and much back-tracking. This is behaviour characteristic of ants when they are disorientated or lost (MacGregor, 1948; Carthy, 1951).

In the maze, their tortuous tracks at first keep them in and later repeatedly bring them back into the entrance alley of the T. Later their tracks gradually extend farther along the arms of the T, but even after entering these they often do not reach the end of an alley before turning back. During these early stages ants tend to walk on the walls and the roof of the maze and try to find a way out in edges and corners. At a later stage in the learning process, about the fourth or fifth trial, the amount of back-tracking shown in the entrance alley decreases, as does the tendency to walk on the walls. The ants still show, however, a considerable amount of back-tracking in the arms of the T, often repeatedly passing from one arm to the other without travelling to the end of either. As training proceeds further the tracks into the blind alley progressively shorten and the extent of tracks into the exit alley increases, so that the ant more often reaches the end of the exit alley. Even so, it will often turn back without entering the exit tube. When the ants have finally reached a good level of performance, they show little hesitation in running through the maze, and their tracks start to curve towards the positive alley while the ant is still in the entrance alley.

If during a period of good performance the ant is disturbed in some way, by, for example, an unexpected vibration of the maze, or is hyperexcited owing to mis-handling during transfer from the nest, the convoluted tortuous tracks recur, and the number of errors then increases. Similarly, if the ant is in an aggressive mood, it can often be observed running backwards and forwards in the maze, biting any edge of the linings, which it can take hold of, and generally apparently 'looking for something to attack'. When such disturbance occurs, it is obviously incorrect to say that the learned behaviour has been forgotten; it merely implies that the ant is no longer motivated to leave the maze, and so does not return to the nest. Fortunately, it was usually immediately clear when the ants were disturbed in this way, and if they could be caught in the entrance alley they were temporarily removed from the maze, before errors had occurred. They were then reintroduced later.

Three measures of learning were taken in preliminary experiments. These were the

length of track, time spent in maze and number of errors. It was found that the time spent in the maze was not a very good measure of progress of learning, since some ants would sometimes make a very slow but error free run, while others (or the same ant on another occasion) might make a fast run with many errors. The length of track was good measure, but it was difficult to obtain accurately, and was also, of course, laborious to take, particularly in the early trials. Subsequently, the error score only was used, since this was not only the easiest to record, but also reflects the track length, which itself is a measure of disorientation. An error was scored whenever an ant either completely entered the blind alley, or completely entered the exit alley, but turned back and left again by the entrance tube. This type of score gives a good measure of the ant's progressive learning, and also a good indication of disorientation if ants are disturbed by a change in the stimuli after learning.

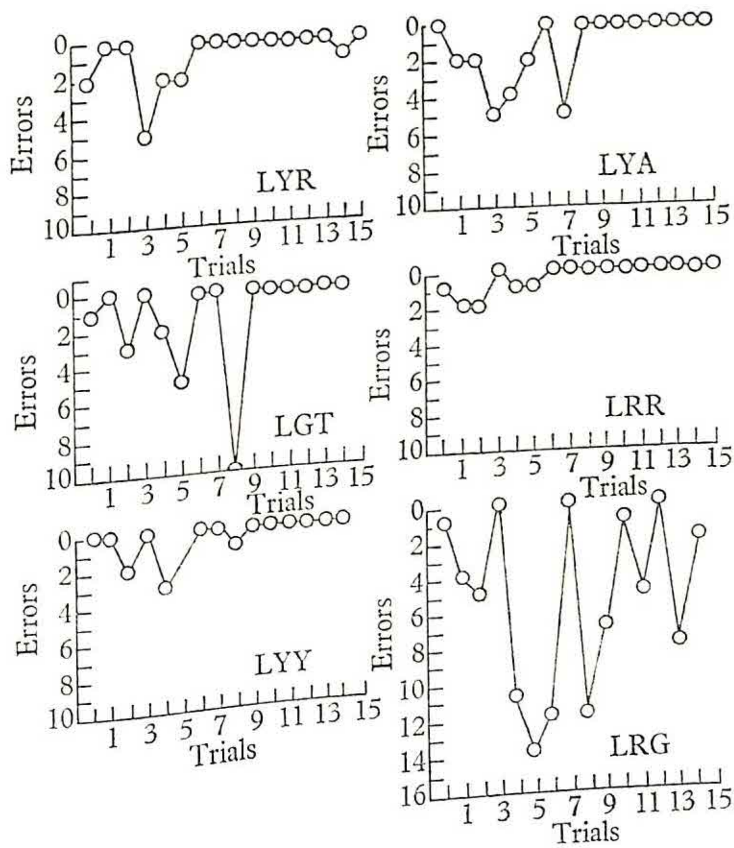


Fig. 2. Learning graphs of five individual ants trained in vertical versus horizontal stripe discrimination. The letters refer to individual reference codes.

In Fig. 2 some typical learning graphs are shown. It will be clear from these that the T-maze for ants cannot be simply regarded as a two-choice situation in which only one error can be made per trial—as, for example, in a Lashley jumping stand, or a T-maze with swing doors used with rats. Ants seldom show a sequence of trials with only one or zero errors. This is because their back-tracking behaviour increases the number of errors in any trial when the ant is disorientated. It is therefore necessary to decide what constitutes a suitable criterion of learning.

The mean error score itself is not a good measure, owing to the large number of errors which can occur on a single trial. For example, if in a series of ten trials nine are without error but one has ten errors (which is a common result) this gives a mean score of one error per trial: which could also be obtained by an ant making no correct runs and actually scoring one error per trial. A better measure is to count the number of trials on which errors occur, irrespective of the number of errors.

As noted, the maze cannot be regarded as a simple two-choice situation. Errors may be due both to an ant entering the blind alley *and/or* to an ant entering the correct alley but turning back and failing to leave through the exit on that occasion. The ant has therefore *at least* three alternatives—entering the blind alley, entering but turning back from the exit alley, or entering and actually leaving by the exit alley. However, the precise probability can only be determined empirically by studying the ants' actual behaviour.

In a sample of forty ants the number of trials without errors on the first trial was 8. This is small because of the high level of back-tracking, characteristic of unfamiliar surroundings. If, however, the ant learned the maze as such, but was unable to identify the exit alley, the number of error free trials occurring by chance might be larger. One way of obtaining a measure of such chance occurrences would be to count the number of error free trials shown by ants when set an insoluble problem after a stage at which they would normally have been performing well on a soluble problem. As described later ants seem unable to learn diagonal *v.* diagonal stripe discrimination. In a group of fifteen ants set this problem, on the 20th trial five made an error free run. If five out of fifteen is accepted as the probability of making one error free trial by chance, the probability of five consecutive error free trials is therefore $0{\cdot}33^5 = 0{\cdot}004$. This criterion was therefore accepted.

In any single nest there were usually ten individual ants used for training. Of these two or three usually died before the end of the experiment and another one or two individuals failed to learn within the limits of the experiment (about thirty trials). There is considerable variation in individual performance between individuals who did learn. Thus some ants may learn in two or three trials while others take twenty-five or twenty-six. If one wishes to compare the discriminability of different patterns by using the number of trials taken to learn to distinguish between them, some account must be taken of individual variability. Unfortunately a considerable difference was found to exist between the performance of ants coming from different nests and this is thought to be due to differences in the nest conditions, rather than to genuine inter-individual differences. For the purpose of comparison it is therefore necessary to use ants which come from the same nest at the same time on both problems. In the experiments reported here this conclusion was not reached until after the end of the experiments, and ants which learned different problems came from different nests. It is therefore not possible to conclude from the results reported here anything about the relative ease of discrimination of the different patterns used.

Another difficulty which arises in these experiments is the choice of a criterion to decide whether ants are disturbed by a change in experimental conditions. In the experiments on generalization for example, ants are trained using stripes of particular dimensions and after they have learned, tested with stripes of different dimensions. It is then necessary to know whether the ants were disturbed by a change

in the stripes or whether they generalized from one pattern to another. When ants are disturbed by a change in conditions, they show a recurrence of back-tracking behaviour which leads to a recurrence of errors. However, it is necessary to know for a control how consistent an ant's performance is under unchanged conditions. The learning curves for some individuals which have been overtrained are shown in Fig. 3. It is clear from the graphs that ants are very consistent in their performance under constant conditions. In twelve ants overtrained on a vertical versus horizontal stripe discrimination it was found that after the initial five error free trials, mistakes occurred only at a mean level of 1·6 trials with errors out of ten trials. Thus in any experiment the control group show a sufficiently reliable and consistent performance for comparison with the experimental group.

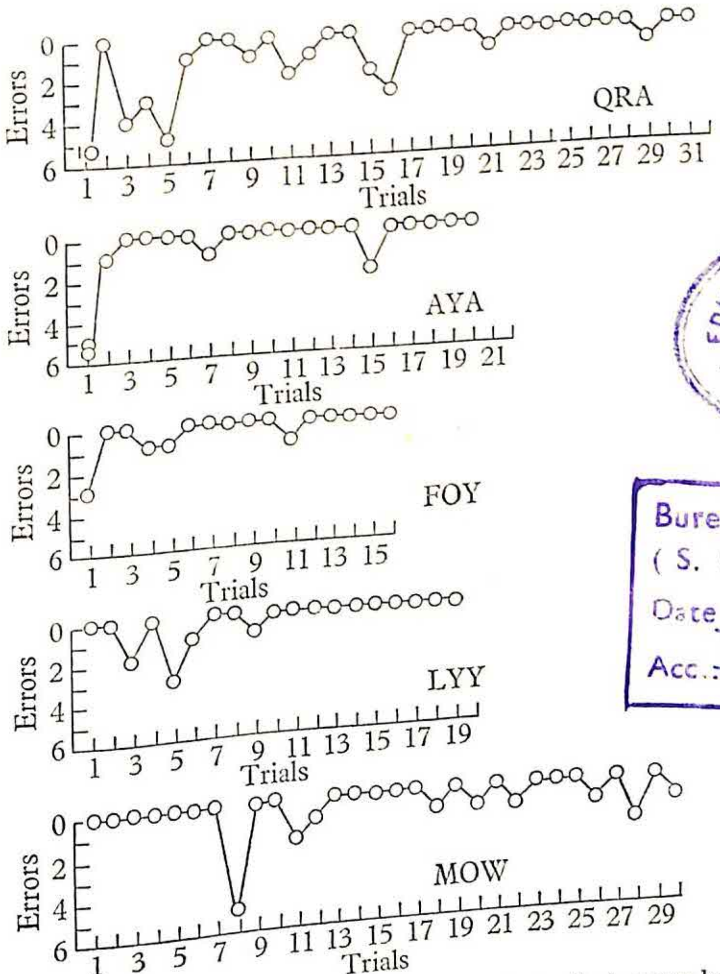


Fig. 3. Learning graphs of five individual ants overtrained in vertical versus horizontal stripe discrimination. The lower graph shows the deterioration which often occurs with prolonged training.

A further complication which is involved in this situation is that if training and testing are continued for more than a total of thirty trials, the performance of an ant often deteriorates. Thus it is not feasible to continue the experiments for a large number of trials. Further, if an ant is disturbed by a change in conditions but is repeatedly exposed to these changed conditions, it will presumably start to learn the

new situation. Since learning may occur in a small number of trials it is therefore necessary to select for analysis a sufficiently small number of trials after conditions have changed to reduce or exclude the possibility of improvement in performance due to learning. It was found empirically that six trials following the original learning (i.e. a total of eleven trials from the initial reaching of criterion) provided a suitable number for analysis. Using a median test the number of trials showing errors in the six trials following a change of conditions were compared with the six trials following original learning in a control group. A one-tail test was used with the hypothesis that a change in the conditions should disturb the ant's performance. This is a crude test which does not distinguish degrees of disturbance.

An attractive experimental procedure which was tried initially was to mix test trials (for controls) with training trials using the original conditions: unfortunately disturbance in the test trials also affected performance on the training trials, and often abnormal behaviour developed in the maze. The procedure was therefore abandoned, and separate groups of ants used as controls.

RESULTS

Vertical versus horizontal stripes

If insect pattern vision is based on the quantity of flicker produced by a figure when its image moves over the compound eye, then a series of vertical stripes should be easily discriminated from a series of similar horizontal stripes by an ant running on a horizontal surface. If, however, only the length of contour is important such stripes should be indistinguishable. The possibilities were tested as described below.

The patterns which the ants were required to distinguish from each other in this series of experiments were composed of parallel red and white stipes of approximately 2 mm. width and separation. This width is well within the limits of the ants' visual acuity—partly because the ant is close to the stripe, which therefore subtends a large angle. The ommatidial angle of the wood ant is about 4–5° in the horizontal plane (from Werringloer, 1932). The stripes were red rather than black merely because this colour was supplied commercially as 'Zipatone', and black stripes of similar dimensions were not then commercially obtainable. The stripes did not have very sharp or straight edges when examined at high magnification, but appear reasonably clear to a human observer. Insects are known to be relatively insensitive to red, and it is probable that the red stripes do appear almost black to the ants. In some pilot experiments, hand drawn and photographically reduced black and white stripes were substituted after the ants had learned the red discrimination; no difference in behaviour was observed.

In one alley, and on the adjoining side of the entrance alley, the walls were lined with horizontal stripes: the other alley and other side of the entrance alley were lined with vertical stripes. Training proceeded as already described. Half the population of ants were trained with vertical stripes positive, and half with vertical stripes negative. Typical results for several individuals are shown in Figs. 2 and 3. The writer has used this training situation as standard for other purposes, and many more ants have learnt this discrimination than are shown in these figures.

It will be clear from Fig. 2 that, with the exception of ant LRG, all the individuals successfully learnt the discrimination. Ant LRG is included in order to indicate the sort of performance shown by an ant which does not learn the maze. The positive results support the suggestion that flicker is an important feature of form perception, although other hypotheses could predict the same result.

When the early performance of the ant is at a high level there is usually a falling off before learning. Similarly, a deterioration in performance occurs towards the end of a long training period. A qualitative examination of all the results for learning a T-maze by ants suggests that learning is often most rapid in ants which show an initially poor performance, rather than in those which show an initially good result. Possibly this may be due to the fact that the latter group of ants have less experience of the blind alley.

Taking as a criterion for learning the first of a succession of five consecutive error free trials, and omitting the very few ants which did not reach this criterion within thirty trials, the mean number of trials to reach the level is 14·5 with a S.D. = 3·6, $N = 22$.

Perception of diagonal stripes

If the important cue in form vision is the quantity of flicker caused by the pattern, ants should be unable to distinguish identical rectangles tilted at 45° to the right and left of vertical. Such a discrimination is also diffcult for many other animals (Sutherland, 1962).

In these experiments therefore the ants were required to discriminate between the direction of slope of otherwise similar patterns. The 2 mm. red and white stripes described in the previous section were tilted at an angle of 45° to the horizontal.

Care was taken to arrange the stripes so that the direction of slope is different relative to the ant, rather than to the external human observer. If the stripes are made to slope up to the left and right in the two arms of the T these look different to the experimenter. However, to an ant running past the stripes such an arrangement presents precisely the same stimulus in the two arms and the discrimination cannot be learned. The stripes were therefore arranged to slope downwards toward the end of one alley, and upwards towards the end of the other.

Two nests of ants were initially trained on this discrimination, and some of their results are shown in Fig. 4. It is clear from these graphs that no individual succeeded in reaching the criterion for learning adopted in the previous section. Some individuals did show a plateau in their performance, making errors fairly randomly, between 0 and 3 per trial. This is the sort of standard which might be expected if individuals were able to respond to the T-maze as such, but not to the stimuli which could lead to correct performance. Other individuals shown varied considerably more in their performance. Trials were discontinued either because the ants died or because after thirty trials the ant's behaviour was becoming abnormal.

In order to guard against the possibility that some adverse characteristic of the nests was responsible for this failure in learning, the experiment was repeated with another nest, other individuals of which had previously learned another discrimination. The results of this experiment are shown in the lower part of Fig. 4. It was clear again that none of these individuals had learned within thirty trials. After

this period three of them were trained with vertical versus horizontal stripes, in place of the diagonal patterns, which they easily learned. This is indicated by the solid circles on the graph. In the two experiments a total of seventeen ants failed to reach the criterion for learning.

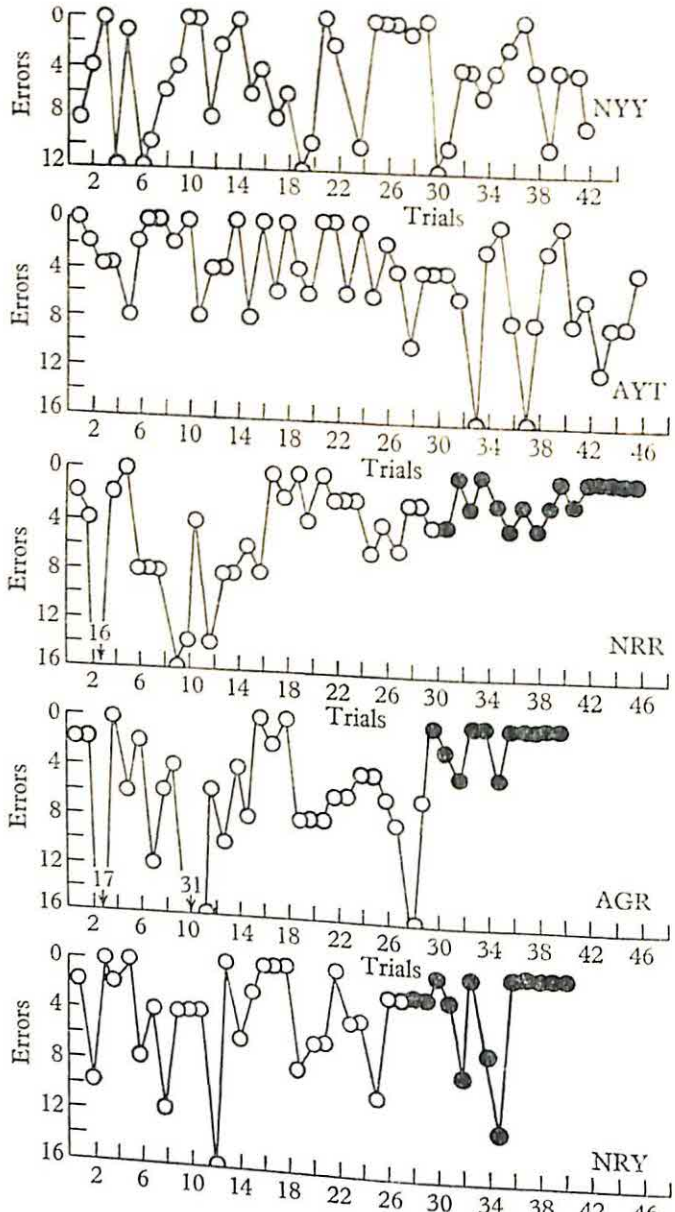


Fig. 4. Learning graphs for ant trained in diagonal versus diagonal stripe discrimination. In the lower three graphs the ants were subsequently trained in vertical versus horizontal stripe discrimination (solid circles).

It seems therefore that ants are unable to discriminate between diagonal stripes of opposite slope, at least within the time limits which have proved adequate for learning other discriminations. The failure to learn this pattern also implies that the ants were unable to learn to alternate during the experimental period. This is an important control for subsequent experiments, when a change in experimental conditions failed to disturb the ants (see later).

Vertical versus diagonal stripe discrimination

If the flicker involved in form vision is caused by movements of contours over the retina then ants should also be unable to distinguish a vertical from a tilted contour. As Wallace (1960) has pointed out, single contours at different orientations, but of the same vertical extent, will sweep over equal areas of the eye as an insect moves past (parallelograms on the same base and between the same parallels are equal in area). However, in repetitive patterns such as stripes the horizontal width of the stripes will also be important—the narrower the stripe the higher the frequency of flicker as the ant runs past.

In these experiments therefore ants were trained to discriminate between vertical and diagonally striped patterns. Two conditions were used: in one of these the two patterns were composed of the same stripes, one vertical and the other tilted at 45°. In this condition the absolute widths of the stripes and their spacing are equal for both patterns, but the widths of the stripes and spaces on a horizontal axis are greater in the diagonals. In the other condition the diagonal stripes were of such a width and separation that their measurements on the horizontal dimension were in fact equal to those for the vertical pattern. These diagonal stripes were therefore narrower on their transverse axis. The stripes were prepared in both cases by drawing large black and white patterns manually, and reducing them photographically. As before, the basic dimensions of the vertical stripes were 2 mm. stripe width and separation.

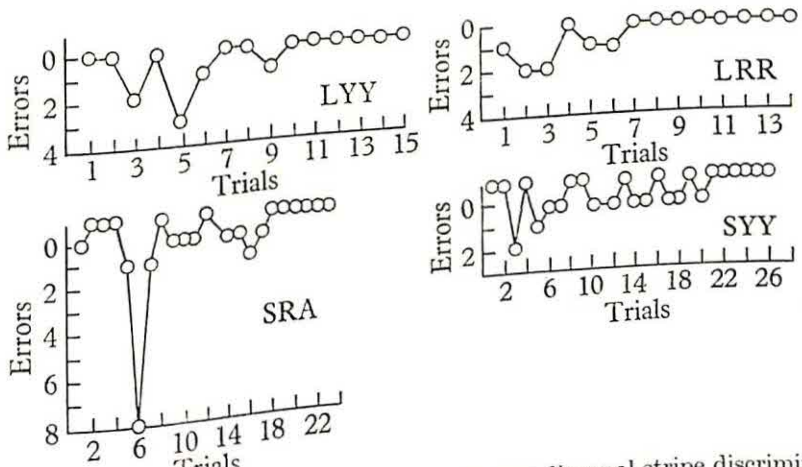


Fig. 5. Learning graphs for ants trained in vertical versus diagonal stripe discrimination. The upper two graphs were obtained using stripes of equal horizontal width, and the lower two with stripes of equal absolute width.

The results of these experiments are shown in Fig. 5. Six ants were trained under the first condition and five ants under the second condition. The vertical stripes were positive in three individuals under the first condition and three individuals under the second condition. It will be clear from the graphs that these discriminations were adequately learned. The mean number of trials to reach the learning criterion were 14 and 9·6 respectively, but although this difference is suggestive, the ants came from different nests at different seasons, and no conclusions can be drawn as to the relative difficulty of the discriminations.

Discrimination of vertical stripes of different widths

If an ant is running close to an object for a given distance the area swept out by a single contour and hence the number of stimulated ommatidia will be greater than when the object is distant. If the ant cannot allow for the effects of distance it might confuse similar objects of different sizes at different distances, or the same sized objects at different distances, since they might produce the same total flicker in the same time.

As Wallace (1960) has shown, one of the cues used in distance estimation is the rate or extent of movement of contours over the eye when the insect performs a voluntary scanning movement. This might lead to some difficulty for ants in the maze situation since the same sort of stimulus appears to be used for the pattern discrimination itself.

In these experiments therefore the ants were trained to discriminate between vertical striped patterns in which the ratio of the widths of the stripes was 2:1. The striped patterns were again prepared by drawing and photographic reduction. The patterns were of equal stripe width and separation and the three patterns used were stripes, 1, 2 and 4 mm. in width respectively.

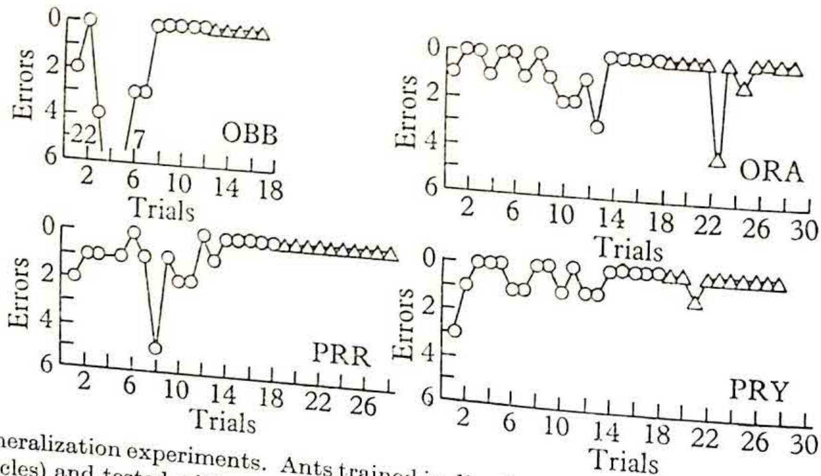


Fig. 6. Generalization experiments. Ants trained in discrimination of vertical stripes of different widths (circles) and tested with stripes of different sizes but the same ratios (triangles). Upper graphs trained on stripes sizes 4:2 and tested on 2:1. Lower graphs trained on stripes sizes 2:1 and tested on 4:2.

In one group of experiments the ants were trained to discriminate between 2 and 1 mm. stripes, the 2 mm. stripes being positive. After learning, the ants were presented with a choice between stripes of width 4 and 2 mm., where the 4 mm. stripes were positive (i.e. led to the nest). In the other conditions ants were trained initially to discriminate between stripes of 2 and 4 mm. width, again with the 2 mm. stripes positive. After learning they were presented with a choice between 1 and 2 mm. width stripes in which the 1 mm. pattern was positive and led to the nest.

The experiments were planned in this way to see if ants *could* respond to relative sizes of the stripes. In each case the previously positive stimulus became negative. The comparable experiment to check if they were using measures of absolute size was not done in view of the positive results for the first experiments. Some results

of this experiment are shown in Fig. 6. Five ants were trained in condition one and four ants in condition two. There was no significant difference between the two conditions. The mean number of trials on which errors are shown during the first six trials after the change of stimulus is 1·3, $N = 8$. A median test in which these results are compared with the control group in which continuous training was given with no change of stimulus after the five consecutive error free trials gave $P = 0{\cdot}16$, which means that the performance is not significantly different from the control group.

There is thus some evidence that the ants were able to respond to the relative size of stripes rather than their absolute dimensions. In view of the fact that insects are not known to possess a size constancy mechanism, it is relevant to note that in these experiments, particularly during the middle period of training, ants were frequently observed to run close to one wall of the entrance alley and then to cross over and run close to the opposite wall. This behaviour suggested that the ants were successively sampling the two striped patterns from their near vicinity. The ants seemed to be comparing the patterns in time rather than space. If this is so, then obviously a size-constancy mechanism would not be necessary to enable the ants to distinguish between the stripes since they are being compared successively from similar short distances.

General discussion

The purpose of the experiments described above was to devise a method for training ants in visual discriminations, and to use the method to study insect pattern vision. It is not clear whether the ability of ants to orientate successfully through a maze by means of patterns on the walls should strictly be regarded as showing form vision or not. The subsequent discussion should, therefore, be considered with this caution in mind, since the generality of the findings to the more usual type of form perception situation is not known.

Previous work on form perception in bees (Hertz, 1929*a*, *b*, 1931, 1933, 1935) has suggested that the most important feature of the stimulus by which they detect flowers ist he amount of contour in the figure. Wolf & Wolf (1936) have suggested that it is merely the total quantity, or perhaps the frequency, of flicker produced in the ommatidia as the image moves over the compound eye which provides the effective stimulus.

Clearly the total flicker produced by a stimulus in a given time as it is scanned by the compound eye is the sum of the light/dark, and dark/light changes in the individual ommatidia. This will depend on the speed of movement, the separation of the contours, the total number of contours and the total length of the contours.

It is not quite clear, as Wallace (1958) points out, whether some factor other than total flicker, such as the relative orientations of the contours to each other and to the figure as a whole, might not be important. In beetles, Hassenstein (1958) has shown that figure-ground discrimination could be carried out by the same mechanism used in movement perception. In this mechanism the relevant integrations are made almost entirely by groups of two or three adjacent ommatidia, independently of their locus. The effects of the stimulus falling upon other, more widely separated, parts of the eye are merely summated subsequent to their integration. If this same mechanism

is used for *identification* of a figure, as well as for distinguishing it from its background, this would support the suggestion of Wolf that form perception is done in terms of quantity or frequency of flicker produced by the patterns. The 'figural quality' of a pattern might depend merely on the distribution of different contours relative to each other and to the ommatidia: contours falling on adjacent ommatidia would interact to produce greater effects than those falling on more widely separated ommatidia. Further, since the ommatidial angle may vary along different axes (e.g. horizontal and vertical) the relationship of the contours to these axes will obviously be important.

Of the stimuli used to guide the ants through the maze in the present experiments, one would anticipate a ready discrimination between vertical and horizontal stripes. Ants, unlike bees, possess only very slight binocular overlap in their visual fields. In the entrance alley, therefore, the different patterns on the two walls can stimulate the two eyes almost separately. Although the total length of contour is the same one would expect the vertical stripes to produce maximum flicker in one eye as the ant ran past them, while the horizontal stripes would produce a little flicker only as the ant moved slightly in the vertical plane. The discrimination involved could therefore be between the pattern producing considerable flicker and the pattern producing relatively little flicker. As predicted the distinction was easily learned.

In the case of the diagonal stripes, one would predict that ants could not learn to discriminate in terms of the direction of slope. The total frequency or quantity of flicker produced by the same patterns of opposite slope would in fact be the same. If, as Hassenstein has shown, movement is reacted to only as a vector by insects, then the direction of slope of the lines will not be perceived. The prediction is confirmed since the ants did not learn this discrimination.

The experiments on discrimination between vertical and diagonal stripes were also done to investigate the ant's ability to distinguish different slopes of stripes in terms of the possible flicker they produce. When contours of different lengths but between the same horizontal parallels move over the same horizontal distance they sweep over areas which are equal (Wallace, 1959), i.e. parallelograms on the same base and lying between the same pair of parallels are equal in area. This means therefore that a single vertical or sloping contour of the same vertical extent will stimulate the same total number of ommatidia when the ant moves past them for the same distance. However, in the two experimental conditions used the actual rate of flicker will depend also upon the horizontal widths of the stripes, since the ant is moving past them horizontally. In one case these were equal, and in the other the horizontal extents of the sloping stripes were greater. In the latter case the rate of flicker produced by the diagonal stripe will be less than that given by vertical stripes, and in the former the same. One would predict that only stripes of different horizontal widths would be discriminated. The experiments, however, show that both these discriminations can be made, and some other factor besides rate of flicker produced by horizontal movements must be important.

At least three further variables might be involved. One of these is the area of the individual stripes. In the case of stripes of equal horizontal width, the area is the same for both verticals and diagonals, and this factor is therefore eliminated in these conditions. A second variable might be the ability of the ant to scan in vertical as

well as horizontal directions. If, for example, it moved its head up and down vertically while running, then its vertical scanning movement will cause more flicker from the diagonal pattern than the vertical pattern. Again, this would provide a cue for distinguishing the two patterns. A difference in vertical and horizontal ommatidial angles might also make contours of different slope differently effective as stimuli, for any one direction of movement. A further factor might be the rate of change in intensity of light to dark and vice versa: this will be greater for a vertical contour than a diagonal one for each horizontally moving ommatidium.

This series of experiments throws a little light upon the mechanism of form perception in insects. They are explicable in terms of previous theories that this mechanism is primarily based upon flicker, but since the movement of the ant relative to the patterns was not precisely controlled, little can be said of the type of scanning necessary to produce such flicker. Perhaps a more interesting approach might be to use patterns which are themselves flickering or moving. This has not been attempted.

The experiments on transfer between vertical stripes of different dimensions were done to test whether insects could allow for the effects of distance and respond to relative proportions of patterns. The stripes used for discrimination were placed on all the walls of the maze, and both patterns are therefore present in the entrance alley. It was clear from the ant's movements and orientation that, after it has learned, its choice is made while in the entrance alley, before proceeding to the junction of the T. Since the ant can move freely within the entrance alley, it is sometimes closer to one wall than the other, and the angles subtended by the stripes at the ant will vary with its distance from them. This raises the problem of how the ant can allow for apparent size differences in the stripes, when these apparent differences depend upon its distance from them.

Only two types of distance perception have been described in insects. In using one of these, the dragon fly nymph (Baldus, 1926) seizes its prey by projecting its raptorial mask when the image of the prey falls on particular groups of ommatidia on the front of its compound eyes. When these ommatidia are stimulated, the prey is in fact at the correct distance for the mask (of fixed size) to strike it. This does not seem to be a general type of mechanism for distance perception. The second method for distance estimation has been demonstrated by Wallace (1959) for the locust. This insect sways its head from side to side while peering at objects before jumping on to them. The distance of the object is apparently judged by the movement of a contour over the retina. It does not seem to require parallax effects, and can be used to control the accuracy of the jump towards a single object against a homogeneous background. Such a procedure could obviously be of general importance in distance perception by insects.

A similar scanning effect might be produced by ants as they run past a striped pattern. The ant would not need to measure its actual speed, but merely to compare the rate of movement of the contours over the eyes looking towards the two patterns. If the ant is also comparing the size of the stripes in terms of the rate of flicker produced as it moves relative to them, it would be necessary to have a separate mechanism for measuring the velocity of contours and the flicker produced by whole figures. From Hassenstein's (1958) work, it seems that the perception of moving figures *and* their velocity may be carried out by the same mechanism. If this is the case, one

would not expect the ant to be able to allow for the effect of distance on apparent size.

However, the ants did transpose between patterns of the same relative, but different absolute, proportions. Observations of the behaviour of the ant in the maze during learning these discriminations suggest a mechanism whereby this might be accomplished. The ants were very frequently observed when in the entrance alley to run close to one wall, palpating it with their antennae, and then, after a short distance, to cross to the other side of the alley and repeat the process on the other wall. The behaviour gave the observer a distinct impression that the ants were comparing the stripes not simultaneously in space but successively in time, the successive comparisons being done from the same distance in each case. Similarly, ants would often run close to the end wall of the T where both patterns are present, turning back on their tracks if necessary for a correct solution. Under such conditions the ant could obviously use a temporal comparison of the frequency of flicker produced by the two equally close patterns, and thus avoid any difficulty there might be in allowing for changes in apparent size with distance. If this interpretation is correct it would support the suggestion that ants have no mechanism for size constancy at different distances.

It is perhaps relevant to consider the results obtained with ants in relation to those for vertebrates. One significant point seems to be that under the experimental conditions used, ants will learn visual discriminations of a simple sort very rapidly. Their ability to learn to discriminate between vertical and horizontal stripes for example, compares favourably with that of the cat, the rat, the monkey and the octopus (Sutherland, 1962). It is also interesting that, like rats (Lashley, 1938) and octopuses (Sutherland, 1957, 1958), ants have difficulty in discriminating between diagonal stripes of opposite slope, although the mechanisms leading to this difficulty are probably different in all three animals. If ants are reacting to the quantities of flicker produced by scanning an object it may well be that they do not possess a mechanism which distinguishes figures from their background. Although the experiments reported here establish that some form of pattern discrimination exists, it may well be that we are studying the 'flow patterns' produced by moving contours, and used by the ant to steer itself, rather than a mechanism for identifying figures as such.

The research reported in this paper has been sponsored in part by the Air Force Research Division of the Air Research Development Command, United States Air Force, through its European Office, under Contract Nr. AF 61(052)-420, AF 61(052)-114. I also wish to express my gratitude to the Royal Society and the Nuffield Foundation for their generous support and to Mrs J. Wallace for technical assistance.

References

BALDUS, K. (1926). Experimentelle Untersuchungen über die Entfernungslokalisation der Libellen (*Aeschna cynea*). *Z. vergl. Physiol.* **3**, 475–505.

CARTHY, J. D. (1951). The orientation of two allied species of British ants. I. *Behaviour*, **3**, 275–303.

ELLIS, P. & PEARCE, A. (1962). Group formation in locust hoppers. *Anim. Behav.* **10**, 305–18.

HASSENSTEIN, H. (1958). Über die Wahrnehmung der Bewegung von Figuren und unregelmässigen Helligkeitsmustern. *Z. vergl. Physiol.* **40**, 556–92.

HERTZ, M. (1929*a*, *b*, 1931). Die Organisation des optischen Feldes bei der Biene, I, II und III. *Z. vergl. Physiol.* **8**, 693–748; **11**, 107–45; **14**, 629–74.

HERTZ, M. (1933). Über figurale Intensitäten und Qualitäten in der optischen Wahrnehmung der Biene. *Biol. Zbl.*, **53**, 10–40.

HERTZ, M. (1935). Die Untersuchungen über den Formensinn der Honigbiene. *Naturwissenschaften*, **23**, 618–24.

HERTZ, M. (1937). Beitrag zum Farbensinn und Formensinn der Biene. *Z. vergl. Physiol.* **24**, 413–42.

LASHLEY, K. (1938). The mechanism of vision. *J. Gen. Psychol.* **18**, 123–93.

MACGREGOR, E. C. (1948). Odour as a basis for orientated movements in ants. *Behaviour*, **1**, 267–97.

SCHNEIRLA, T. C. (1941). Studies on ant learning. *J. Comp. Psychol.* **32**, 41–82.

SCHNEIRLA, T. C. (1953). Basic problems in the nature of insect behavior. In *Insect Physiology*. Ed. Roeder, K. D. New York: Wiley.

SUTHERLAND, N. S. (1957). Visual discrimination of orientation by octopus. *Brit. J. Psychol.* **48**, 55–71.

SUTHERLAND, N. S. (1958). Visual discrimination of the orientation of rectangles by the octopus. *J. Comp. Physiol. Psychol.* **51**, 452–8.

SUTHERLAND, N. S. (1962). Shape discrimination by animals. *Monograph Exp. Psychol. Soc.* **1**.

VOWLES, D. M. (1955). The foraging of ants. *Brit. J. Anim. Behav.* **3**, 1–13.

VOWLES, D. M. (1958). The perceptual world of ants. *Anim. Behav.* **6**, 115–16.

WALLACE, G. K. (1958). Some experiments on form perception in nymphs of the desert locust. *J. Expt. Biol.* **35**, 765–75.

WALLACE, G. K. (1959). Visual scanning in the desert locust. *J. Exp. Biol.* **36**, 512–25.

WALLACE, G. K. (1960). Visual perception and orientation in the desert locust. *Brit. J. Anim. Behav.* **8**, 231.

WERRINGLOER, A. (1932). Die Sehorgane und Sehzentren der Dorylinen nebst Untersuchungen über die Facettenaugen der Formiciden. *Z. wiss. Zool.* **141**, 432–520.

WOLF, E. & ZERRAHN-WOLF, G. (1936). Flicker and the reactions of bees to flowers. *J. Gen. Physiol.* **20**, 511–18.

(*Manuscript received 2 September* 1963)

Brit. J. Psychol (1965), 56, 1, *pp.* 33–44
Printed in Great Britain

RECENCY IN TWO-CHOICE GUESSING TASKS

By H. C. A. DALE
Medical Research Council Applied Psychology Research Unit, Cambridge

It has been asserted that negative recency effects are found only in simple tasks where the alternatives remain constant from trial to trial. The experiments reported show that, with alternatives varying from trial to trial under conditions where a simple, obvious rule appears to exist relating the correct alternative to a clue, positive recency is found. But, under conditions where the choice lies between varying alternatives which come from two simple categories, negative recency predominates.

If in a series of binary choices the subject's guess tends to correspond with the previous event, he is said to exhibit 'positive recency'. On the other hand, if he tends to predict the event which did not occur on the previous trial, then he exhibits negative recency.

Negative recency has been observed in a wide range of two-choice situations, both where the random nature of the events is explicitly stated, as when subjects have to predict the fall of a coin, and where no information is available about the way the events are generated (Cohen & Hansel, 1955). Thus it occurs when the experimenter announces that he is going to say 'check' or 'plus' on each trial (Jarvik, 1951) or that he is going to switch on a green or a red light and the subject must guess which it will be (Nicks, 1959). It also appears in many studies of two-choice probabilistic learning as an embarrassment to theories of the kind advanced by Bush and Mosteller, and by Estes (Feldman, 1959; Cohen, 1960, pp. 38–9).

Exceptions have been noticed, however. One of particular interest is reported by Green in which simple adjectives were used as stimulus material (Green, 1958). One adjective would be read out with the instruction that this provided a clue to a 'target' which was another, related adjective. The task was introduced as a guessing game, the subjects having to guess what the target would be. In each case the target was either a synonym or an antonym of the clue. The sequence of events was either S S S A S S S A A A A A or A A A S A A A S S S S S (S standing for synonym, A for antonym). Green found that if three synonyms had been given most subjects would predict a synonym on the fourth trial and vice versa for antonyms. In a control condition, subjects had to guess which of two lamps the experimenter would light at each trial, a green one or a red one. Although the instructions for the two conditions were very similar, the results were quite different. After three 'reds' in the control condition most subjects predicted 'green', thus exhibiting negative recency or, as it is sometimes called, the 'gambler's fallacy'. Green suggests that the essential difference between his experimental and control conditions lies in 'task complexity': '... although synonyms and antonyms do make a binary choice situation, there is still a certain amount of variety possible within these two classes. In the light guessing situation, the subject appears to reject the simple repeated response as too easy, whereas some sort of skill is required to find a related adjective'.

Goodnow (1955) compared behaviour in an insoluble two-choice problem solving task with that in play on a 'one-armed bandit' gambling machine. In the problem

solving task differing sets of geometric figures were used, a key figure being presented together with two target figures, one always being produced by augmenting the key, the other by omitting some parts. Her data revealed negative recency with the gambling task, but neither positive nor negative recency in the problem solving situation. This led her to assert that stimulus variability was a critical differentiating feature: recency effects are found '...where the task is seen by the subject as one involving chance and where the stimulus on a given trial differs from the stimulus on any other trial only in its temporal position in the series'. Thus, whereas Green focuses attention upon similarity between successive responses, Goodnow emphasizes critical features of the stimulus situation. Both conclusions are to some extent subjective. They provide no way of telling what subjects regard as too easy, or as 'tasks involving chance'; so that, although seemingly reasonable, they give no rule for predicting when negative recency is to be expected.

The present investigation took Green's study as a starting point. Its aim was to determine more exactly the nature of the conditions which govern the form of the recency effect. A close examination of Green's two contrasting situations reveals that they differed in a number of ways. With the lights it was immediately clear to the subject that only two alternatives existed; these remained constant from trial to trial and they were ready-named. With the adjectives by contrast, the instructions stated that the target was a related adjective without indicating the form of the relationship, and so the subject had to select his response from an extremely large number of possibilities. Furthermore, he was given no indication that the responses fell into two classes and was presumably unaware that an 'either/or' choice existed at each trial. Then again, the specific target varied from trial to trial and it had to be named by the subject. Finally, there remains a distinction between choosing directly between two kinds of stimuli, and choosing between two kinds of relation, in Green's particular instance between similar and opposite. These are distinctly different aspects of the tasks which could each separately, or in some combination, determine the subject's reaction.

The plan of the investigation was to begin by replicating Green's adjective guessing task and then to modify it until it evoked negative recency. In all the experiments a single pattern of events was employed; this was an extension of Green's sequence to S S S A S S S A A A A A A A A A A A A A.

Experiment I

The primary aim of this experiment was to examine the hypothesis that Green's unusual results using adjective guessing arose because his subjects were given no clear indication that only two kinds of event could occur.

Method

The subjects were divided into three groups, A, B and C. All were given the adjective guessing task, in which a clue led to a target which was either of synonymous or opposite meaning. The condition for Group A replicated Green's. The subjects were given a clue and had to think of a related adjective. Group B were given advance warning that the target would be either a synonym or an antonym. Group C had similar advance warning, and class was emphasized rather than the particular response word.

Printed instructions for Group A read: 'This is a sort of guessing game. I have a list of adjectives which I am going to read out one at a time. Each of these words provides a clue to the target word which is another adjective related to the clue in some way. When I give you the clue adjective, your job is to try to guess what the target adjective is going to be. So write down the clue adjective as soon as I give it, and beside it write down what you think the target adjective is.* When you have done this raise your hand. As soon as I see you have all finished I shall give you the answer and then go on to the next clue. Always try to guess what the target is going to be even if you are not very sure, but don't write it down *after* you have been told what it is, or the whole purpose of the experiment is lost.'

For Group B the following sentence was added at the point marked *, above: 'You will notice that the target has either a similar meaning to the clue or else is its opposite'.

For Group C the foregoing sentence was inserted, and in addition: 'The important thing is to guess which it will be, similar or opposite. This matters more than complete accuracy, and if you cannot think of the actual word just put "similar" or "opposite".'

At the end of the session, after the experimenter had discussed the purpose of the experiment, the subjects were invited to describe briefly in writing the way they had carried out their task.

The experiment was run in two parts in order to counterbalance the material. Table 1 shows the list of words used. Wherever a synonym was the correct target in Part 1, an antonym was correct in Part 2. In this way contamination of the results by a preference for particular target words was avoided. Subjects from all three groups were tested simultaneously. Two separate batches were employed; one in Part 1 the other in Part 2. The subjects were twenty-eight male and eleven female volunteers from the G.P.O. Training School at Bletchley, Bucks. Their ages ranged from 16 to 50.

Table 1. *The lists employed in Experiment I*

	Clues	Targets In Part 1	Targets In Part 2		Clues	Targets In Part 1	Targets In Part 2
1	Quick	Slow	Fast	11	Expensive	Dear	Cheap
2	Wide	Narrow	Broad	12	Old	Ancient	Young
3	Gay	Sad	Happy	13	Nice	Pleasant	Nasty
4	Feeble	Weak	Strong	14	Tiny	Small	Big
5	Dull	Bright	Dim	15	Hard	Difficult	Easy
6	Wet	Dry	Damp	16	Taut	Tight	Loose
7	Short	Long	Brief	17	Hazy	Cloudy	Clear
8	Bad	Evil	Good	18	Angry	Cross	Pleased
9	Near	Close	Far	19	Rare	Uncommon	Common
10	High	Tall	Low				

Results

The critical feature of the results is the response at trial 4, when the first switch in the sequence occurred. Subjects exhibiting negative recency are correct on this trial. If, on the other hand, they expect the regularity of the first three trials to persist, they will be wrong. The results are given in Table 2 together with 1% binomial sampling limits which provide a guide for judging the significance of deviations from chance levels of correct responding. For Groups A and B there was clear evidence that subjects expected the regularity to continue.

Group A responded in the same way as Green's adjective guessing subjects, which is to be expected since his conditions were being duplicated. The similarity of Group B's behaviour indicates that this cannot be attributed to their ignorance that a binary choice existed. Positive recency occurs even though the subjects have been told that each answer is either a synonym or an antonym of the clue.

Group C gave a much 'weaker' positive recency response at trial 4. Whereas 1 out of 13 in Group A, and 1 out of 13 in Group B were correct, 6 out of 15 were correct in Group C. This difference between Groups A and B, and Group C is statistically

Table 2. *Proportion of subjects correct on each trial in all five experiments*

Exp.	Group	Condition	N	Trial 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	1% binomial confidence limits: Lower	Upper
I	A	Replicating Green	13	0·69	0·85	0·85	0·08	0·46	0·77	0·85	0·08	0·23	0·38	0·77	0·54	0·85	0·92	0·77	0·77	0·85	0·92	1·00	0·16	0·84
	B	Told synonym or antonym	13	0·54	0·77	0·92	0·08	0·77	0·69	0·69	0·15	0·38	0·46	0·54	0·46	0·46	0·77	0·69	0·61	0·69	0·85	0·85	0·16	0·84
	C	Class emphasized	15	0·53	0·53	0·67	0·40	0·40	0·47	0·67	0·26	0·20	0·40	0·33	0·47	0·53	0·67	0·67	0·47	0·67	0·73	0·80	0·18	0·82
II	D	Six alternatives	19	0·42	0·89	1·00	0·05	0·63	0·89	0·95	0·11	0·16	0·32	0·84	0·84	0·89	0·84	0·79	0·68	0·84	0·89	0·84	0·21	0·79
	E	Two alternatives jumbled	19	0·53	0·84	0·84	0·11	0·79	0·84	0·84	0·21	0·26	0·58	0·68	0·79	0·84	0·74	1·00	0·89	0·95	0·95	1·00	0·21	0·79
	F	Two alternatives in lists	19	0·63	0·79	1·00	0·00	0·68	0·84	0·84	0·00	0·11	0·42	0·58	0·63	0·58	0·47	0·68	0·68	0·74	0·63	0·79	0·21	0·79
III	G	Lists headed similar and opposite	20	0·75	0·65	1·00	0·05	0·50	0·75	0·65	0·50	0·30	0·70	0·55	0·75	0·70	0·45	0·75	0·60	0·80	0·75	0·45	0·22	0·78
IV	H	Write *S* or *O*	23	0·39	0·43	0·49	0·52	0·52	0·57	0·49	0·49	0·39	0·39	0·61	0·43	0·43	0·61	0·57	0·43	0·61	0·78	0·78	0·23	0·77
V	J	No clues	56	0·59	0·32	0·23	0·73	0·46	0·50	0·45	0·70	0·43	0·39	0·46	0·32	0·34	0·36	0·37	0·34	0·46	0·48	0·59	0·32	0·68
All		Sequence		0	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1		

significant ($P = 0{\cdot}019$), which suggests that the increased emphasis of this instruction had the effect of making some subjects respond as they would when guessing the fall of a coin. But it must be noted that this statistical test is *post hoc*. A point of interest is that the difference did not occur because subjects in Group C wrote 'similar' or 'opposite' instead of an adjective. Only one subject on one trial made use of this facility.

The response on the second reversal of the series provides additional evidence, although its interpretation is less clear. Up to trial 4 subjects had experienced a simple run and their reaction to this run was registered by their fourth prediction. The reversal at trial 4, however, broke this run, so that the first seven outcomes formed the sequence 0 0 0 1 0 0 0. Subjects could therefore have predicted correctly at the eighth trial either because they employed negative recency, and the fact that this has led to correct predictions earlier in the series might strengthen it, or because they expected the pattern 0 0 0 1 to be repeated. Similarly, incorrect predictions could arise for a variety of reasons. The subjects could have accepted the general regularity, discounting trial 4 as an error. Alternatively, they could have noticed an incipient 0 0 0 1 pattern and assumed that no such pattern should exist in a guessing task, or that this constituted a trap laid by the experimenter.

In Green's experiment, performance at trial 8 was similar to that at trial 4, but the effect of the switch was slightly weaker in both his adjective guessing and his light guessing groups. The effect on Groups A and B of the present experiment was almost as strong as that of the first switch. With Group C the effect was stronger, so that the behaviour of this group became more similar to that of the other two. The proportion correct on this trial was not significantly greater than that in Groups A and B ($P > 0{\cdot}20$). When the responses at both reversals are taken into account, Group C would appear to resemble A and B.

Analysis of the subjects' comments presented problems, largely because of the varying nature of the sequence. Some commented upon only their initial reaction, some upon their confusion when the switches occurred, and others upon the final run. A relatively frequent comment was 'I put down the first word that came into my head'. Such free associating was reported by five subjects of Group A, one of Group B and five of Group C, of whom four, one and two respectively were wrong on trial 4. A number reported using the regularity they detected in a simple and straightforward way. Thus one subject of Group B who was exposed to the A A A S A A A S S. . . sequence wrote: 'To begin with I was looking for opposites. I stayed with that idea for some time but came round eventually to the idea that you had changed and finally wanted similar adjectives'. This subject had persisted with this simple search for regularity despite the reversals in the sequence.

Finding a suitable response word occupied some of them so fully that they were inattentive to the sequential aspect of the task. Thus one subject from Group C wrote: 'More concerned with putting down my first thoughts which most frequently were the opposites, than trying to guess what the target would be'. With some, following the form of relation of the previous trial occurred automatically: 'When I realized an answer was opposite I found myself automatically writing opposites. . .'.

Two members of Group A claimed to have sought some complex rule. One of them

put: 'I was looking for a rule which was not very straightforward. I think I was trying to guess the aim of the experiment in order to find the rule'. But despite this apparent claim to attempt to out-guess the experimenter this subject produced in response to the sequence beginning A A A S: A A A A A A A A A A S A S S S S S S; thus he was wrong on both trials 4 and 8, and he eventually changed and followed the final run of synonyms. The other subject mentioned behaved similarly. The sophistication indicated in their comments was not matched by their behaviour.

Looking for patterns in the sequence was reported by a number of subjects: two of Group A, five of Group B, and two of Group C. One or two indicated guessing behaviour of the sort that is associated with the gambler's fallacy. Thus one subject of Group A stated: 'I was thinking on the lines that you were going from like to opposite, but not knowing when decided to stick to opposites'. Altogether two subjects of Group A, none of Group B, and two of Group C made reports of this kind. Similar comments have been reported by Cohen (1960, p. 36).

When the written comments and the choices are considered together the general picture which emerges is that the subjects were preoccupied with the problem of generating a suitable response word. The correct answers gave a pattern which provided direction, and there is some indication that this was followed unwittingly. But the fact which must not be overlooked is that when an individual's comments were related to his choices they did not always match.

Experiment II

A significant feature of all the conditions used in Experiment I was that the subject actually had to suggest a target himself. In the typical 2-choice guessing task, however, the specific targets are suggested to the subject, who merely has to pick out the one he thinks will be correct on a particular occasion.

There is some evidence that when subjects have to produce responses they are inclined to repeat responses of similar kinds rather than vary their output. This was noted in the written comments of a number of the subjects of Experiment I. The phenomenon has also been demonstrated in serial reaction time experiments (Bertelson, 1961). It is possible, therefore, that the incorrect predictions that subjects made on the reversals in Experiment I, and also in Green's original investigation, were due to the response set induced by the preceding runs. This possibility was investigated in Experiment II. The general plan was to duplicate the conditions of Experiment I, but to make subjects pick out the target from a number of specific alternatives instead of having to produce a word.

Method

Three groups of subjects were used. Group D was given a list of six possible targets for each clue; a synonym, an antonym, one which alliterated and one which rhymed with it, and two which were quite unrelated. The order of these was jumbled. Group E had two possible targets for each clue, a synonym and an antonym, the order of these being jumbled. Group F also had only synonyms and antonyms as alternatives, but the synonyms were all in the left hand column and the antonyms in the right. The material used for Groups D and E is given in Table 3.

Minor changes in the printed instructions adapted them for the difference in conditions and the different source of subjects. Revised instructions read: 'This is a sort of guessing game. I have a list of adjectives which I am going to read out one at a time. Each of these words provides a

Table 3. *Clues and targets used in Experiment II*

Clue	Targets used for Group D						Targets used for Group E	
1 (Quick)	Empty	Thick	*Fast*	Quaint	Square	Slow	Slow	*Fast*
2 (Wide)	Fried	Wild	*Broad*	Round	Narrow	Light	Narrow	*Broad*
3 (Gay)	*Happy*	Yellow	Costly	Sad	Gallant	Fey	*Happy*	Sad
4 (Feeble)	Pretty	*Strong*	Comfortable	Febrile	Weak	Terrible	*Strong*	Weak
5 (Short)	Funny	Blind	Bought	*Brief*	Shifty	Long	Long	*Brief*
6 (Dull)	Dumpy	*Dim*	Bright	Clean	Hot	Full	*Dim*	Bright
7 (Bad)	*Evil*	Dumb	Bald	Good	Sad	Busy	Good	*Evil*
8 (Wet)	Pet	Peaceful	Weeny	Damp	Royal	*Dry*	*Dry*	Damp
9 (Near)	Fine	*Far*	Grumpy	Neat	Clear	Close	Close	*Far*
10 (High)	Tall	Ugly	*Low*	Handy	Nigh	Noisy	*Low*	Tall
11 (Expensive)	*Cheap*	Extensive	Silly	Manly	Dear	Expert	*Cheap*	Dear
12 (Old)	*Young*	Oily	Jolly	Bold	Dirty	Ancient	*Young*	Ancient
13 (Nice)	New	Pleasant	*Nasty*	Deaf	Cold	Twice	*Nasty*	Pleasant
14 (Tiny)	Musical	Shiny	*Big*	Small	Tiresome	Haughty	Small	*Big*
15 (Hard)	Difficult	*Easy*	Handsome	Marred	Frequent	Natural	*Easy*	Difficult
16 (Taut)	Tapered	Novel	Short	Sly	*Loose*	Tight	Tight	*Loose*
17 (Hazy)	Hearty	Rapid	Lazy	Cloudy	Lost	*Clear*	Cloudy	*Clear*
18 (Angry)	*Pleased*	Anxious	Hungry	Vulgar	Cross	Flat	Cross	*Pleased*
19 (Rare)	Green	Uncommon	Willing	*Common*	Fair	Rancid	*Common*	Uncommon

clue to a target word which is another adjective related to the clue in some way. When I give you each clue adjective, your job is to try to guess what the target adjective is going to be. You will see on the next page that you are given a number of alternatives for each clue and its target will be one of these. Write down each clue adjective as soon as I give it, then pick out the adjective which you think is the target and mark it by putting a ring round it. The first clue has to be written in the first space; its target will be among the words on the first line. The second clue goes in the space on the second line and its target is among the words on that line, and so on. As soon as you have picked out what you think the target will be and have ringed it, look up so that I know you have finished. When you have all finished I will give you the answer and then read out the next clue.'

The experiment was conducted in the same general way as Experiment I. Subjects were tested in two batches. Each batch was split up by assigning roughly equal numbers of subjects to each group. The general procedure was the same for all groups, the only difference being in the printed instructions, so that all three groups could be tested simultaneously. An invigilator checked that the subjects guessed before the answers were given.

The subjects in this experiment were young enlisted men. Written comments were not requested from them.

Results

The results are presented in Table 2. It is clear that under all three conditions the subjects expected the runs to continue on trials 4 and 8. The hypothesis that the tendency to predict continuity of the runs was due to the need to *produce* words rather than to *choose* between targets must therefore be discounted.

Experiment III

In Experiments I and II two factors were varied independently. In this way it was shown that behaviour was only marginally influenced by forewarning subjects that the target would always be either a synonym or an antonym, and was quite unaffected by offering them alternatives from which to choose.

In this third experiment both factors were varied simultaneously; subjects were told that the target would be either a synonym or an antonym and were offered specific alternatives from which to pick.

Method

Only one group of subjects was used, Group G. The subjects were run in two batches, using counterbalanced lists. The materials and general procedure were the same as for Group F of Experiment II, but the lists were headed 'similar' and 'opposite', and minor changes were again made to instructions. The middle section of the instructions for Experiment II was deleted and replaced by: 'When I give you each clue adjective, your job is to try to guess what the target adjective is going to be. You will see on the next page that you are given two alternatives for each clue, one of which has a similar meaning while the other is its opposite. Its target will always be one of these. Write down each clue adjective as soon as I give it, then pick out the adjective which you think is the target and mark it by putting a ring round it. The first clue has to be written in the first space; its target will be among the words on the first line. The second clue goes in the space on the second line and its target is among the words on that line, and so on.'

The subjects were again young enlisted men, but had not participated in the earlier experiment.

Results

The pooled results from both batches are given in Table 2. It can be seen that subjects predicted the continuation of the first run as in the previous experiments. They did not make a similar response to the second run, however. It could be

argued that, while the subjects approached the task prepared to detect and make positive use of regularity, the first reversal led them to reassess the situation. With clearly labelled lists, the reversal provided strong evidence against simple regularity. In Experiment I the subjects had to judge for themselves whether the target at trial 4 was inconsistent with the hypothesis that a run of synonyms (or antonyms) was being used. It is likely under these circumstances that they would discount the reversal because they were not completely sure that it conflicted with the hypothesis. The same could have been true in Experiment II, where, although the target words were provided, the relation was not specified. It is not unreasonable to suppose that this reappraisal would only appear with lists labelled 'similar' and 'opposite'.

As noted, responses at trial 4 were the same as in previous experiments. In the present experiment subjects were told that targets were either synonyms or antonyms. The tasks of Experiment II were more complex to the extent that subjects had had to discover this rule for themselves in order to be able to detect and make use of the regularity of the first three trials by trial 4. Since eliminating this aspect of task complexity makes no difference, no influence on responses at trial 4 can be attributed to it.

Experiment IV

To make the experimental conditions still more similar to those of the usual two-choice guessing task, the requirement that subjects should predict the specific word designated as target at each trial was dropped in this fourth experiment. Instead, subjects were required merely to indicate whether they expected the target to be similar or opposite in meaning to the clue. Thus each response was distinguished only by its position in the series. With this modification the task now resembled Goodnow's gambling task, and alternation might be expected rather than repetition. Goodnow, however, when describing the critical conditions for alternation, added the proviso that subjects react in this way when they see the task as one involving chance, but she provided no rules for determining when this interpretation would be given.

Method

One group of subjects was used, Group H. Two batches were run in a counterbalanced procedure, using the same list of clues and the same lists of targets as in Experiment III. The subjects had a blank sheet of paper; all instructions were given orally. They recorded each clue as it was read out and then put *S* or *O* to indicate whether they predicted that the target would have a similar or an opposite meaning. After each target was given, the experimenter confirmed its relation to the clue by saying 'That was similar' or 'That was opposite'.

The subjects were young enlisted men who had not participated in the earlier experiments.

Results

The results are given in Table 2 and are clearly different from those of the earlier experiments with the exception of Group C in Experiment I. Comparison of the numbers correct on trial 4 shows that significantly more were correct than in Groups A, B, D, E, F and G ($P < 0{\cdot}025$).

Although these results differ from those found previously, they clearly do not represent the pattern Green found for guessing lights, since for this an overwhelming number should have been correct on trial 4. These results show neither positive nor negative recency.

Experiment V

A feature common to all of the conditions used in Experiments I to IV is that the subject had to guess whether the relation linking the clue to the target would be of one kind or the other, similar or opposite. This is yet another way in which Green's adjective guessing task differed from most two-choice guessing situations. In the present experiment this difference was eliminated. The same material was again used, but instead of having to guess the direction of a relation, the subjects had only to decide from which of two lists the target was taken. There were no clues.

Method

One group of subjects, Group I, was used in four batches. The subjects were presented with two lists of possible targets, headed *A* and *B*. Instructions printed above them read: 'This is a sort of guessing game. I have a list of adjectives which I am going to read out one at a time. This list is made up of words from the two lists *A* and *B* which are on this page. The first word will be one of the two top ones. The second will be one of the second pairs and so on. Your task is to guess which word from each pair is on my list. Mark each word you guess by putting a ring round it. Only mark one at a time. As soon as you have made your guess put your pencil down and look up. When I see that you have all finished I will read out the answer. You will then go on to the next pair. Always make a guess even though you may not be sure which it will be.'

When the run was completed, subjects in three of the four batches were asked to write a brief description of the way they had made their choices. As before, the procedures for different batches were counterbalanced. The targets were the same as those used in Experiments III and IV.

The subjects were young enlisted men who had not previously served as subjects.

Results

From Table 2 it can be seen that under these conditions negative recency occurs, just as in the usual two-choice guessing situation. Performance improved suddenly at trials 4 and 8, precisely at the points where it had deteriorated in the conditions employed in Experiments I, II and III.

Written comments about the way they made their predictions had been requested from forty-one of the fifty-six subjects given this condition. Thirty-seven responded to the request. Some made abstruse observations; thus one wrote: 'I try to think of the pleasant things but I found I was putting rings round the unpleasant'. Ten indicated they had sought a left-right sequence. Seven tried to use the meaning of the words as a guide in some way. Another eighteen claimed to have used 'guesswork', 'taken pot luck' or 'picked some from each column'. One thought that the experimenter's intonation provided a clue.

This condition is clearly distinguished from all the earlier ones by the responses at the critical trials, 4 and 8. This final modification has made the adjective guessing task comparable to guessing lights. The crucial factor would appear to be that the subjects had to pick items belonging to one list or the other instead of predicting the relation between a target and a clue.

Discussion

This series of experiments has demonstrated that it is possible to produce both positive and negative recency effects using material in which the specific alternatives available vary from trial to trial. Thus Goodnow's assertion that recency effects are confined to situations in which items can be identified only by their temporal position in a series is false.

The distinction between situations which induce positive recency and those which lead to negative recency is exemplified by conditions G and J. These conditions produced marked differences in behaviour yet were similar in many respects: (*a*) the particular words occurring as targets on each trial were precisely the same; (*b*) the choice was binary in each case and the subject was aware that there were only two possibilities; (*c*) the basis of classifying the possibilities remained constant throughout the run in each task—with clues, the choice always lay between 'similar' and 'opposite', and, without clues, it always lay between 'left' and 'right' lists. Thus the similarity is marked. Yet since the tasks differ in their behavioural consequences, some feature of them must distinguish them phenomenally. There are two related differences between them: the use of clues in G but not in J, and the choice between two kinds of relation in G as opposed to a choice between left-hand and right-hand lists in J.

In order to assess the possible effect of these differences it is necessary to consider the nature of the response patterns which were observed. Although opposite reactions were noticed at the critical trials it would be false to conclude that the response patterns were generally similar in nature but of opposite polarity. It is more likely that the responses were different in nature. In condition G, and others like it, the subjects sought rules; on discovering regularity they inferred that a law was operating and then proceeded to use it. In condition J, on the other hand, a substantial proportion of the subjects indicated that they responded as when gambling, and the response that this group as a whole made to regularity resembles the response to runs in attempting to predict random events (cf. Jarvik, 1951; also Dale, unpublished).

The ambiguous response given in condition H can now be explained. The presence of clues led subjects to abandon any tendency to apply gambling habits; they searched for a rule instead. But, as they had to write *S* or *O* instead of finding a particular target adjective, they sought some characteristic of each clue which would allow them to predict at each individual trial whether *S* or *O* was appropriate for that particular occasion. Since no such distinguishing characteristics existed, their attempts were bound to be fruitless and any hypothesis they tried out was bound to yield choices which were random with regard to the criteria of success employed in the experiment.

References

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quart. J. Exp. Psychol.* **13**, 90–102.

Cohen, J. (1960). *Chance, Skill, and Luck.* London: Penguin Books Ltd.

Cohen, J. & Hansel, C. E. M. (1955). The idea of independence. *Brit. J. Psychol.* **46**, 178–90.

Feldman, J. (1959). On the negative recency hypothesis in the prediction of a series of binary symbols. *Amer. J. Psychol.* **72**, 597–9.

Goodnow, J. J. (1955). Response sequences in a pair of two-choice probability situations. *Amer. J. Psychol.* **68**, 624–30.

Green, R. T. (1958). Factors affecting inductive predictions. *Brit. J. Psychol.* **49**, 97–105.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. Exp. Psychol.* **41**, 291–7.

Nicks, D. C. (1959). Prediction of sequential two-choice decisions from event runs. *J. Exp. Psychol.* **57**, 105–14.

(*Manuscript received* 21 *March* 1964)

Brit. J. Psychol. (1965), 56, 1, pp. 45–51
Printed in Great Britain

ON THE SOLVING OF CODE ITEMS DEMANDING THE USE OF INDIRECT PROCEDURES

By ALISON C. CAMPBELL
Department of Psychology, University of Edinburgh

Subjects in experiments using categorization and matching problems have been found to encounter difficulty with procedures that are indirect, and involve transformations of data. The present study finds this difficulty to extend to compound code items, many subjects not being able to cope with the type of indirect procedure that solution of these items was made to entail. It is suggested that the difficulty stems from a need to work with equivalence classes that are negatively defined and purely verbal, or the fact that the data transformations called for involve a change of logical subject.

Subjects in categorization (concept attainment) experiments markedly prefer direct to indirect tests of the hypotheses they propose (Bruner, Goodnow & Austin, 1956). An hypothesis is proved indirectly through elimination of all rival possibilities. Consider a subject who, faced with two mutually exclusive alternatives X and Y, assumes that alternative X leads to the desired goal. The direct test of this hypothesis is to try X. The indirect test of the hypothesis is to try Y and, according as Y does or does not lead to the goal, to draw an inference as to the tenability of the original hypothesis about X. The preference of subjects for direct tests is found to persist even in circumstances where, unlike the instance just cited, indirect procedures would be more appropriate, economical and efficient. Bruner *et al.* suggest that avoidance of indirect procedures may stem from an appreciation that errors are liable to be made during the 'in the head' transformations of data that indirect procedures entail. They further suggest that avoidance of indirect procedures may well be general to a variety of cognitive activities, and not confined just to categorization. Whether subjects are unable competently to cope with indirect procedures, or merely unwilling to attempt their use when direct—although extremely laborious and inefficient—alternative procedures are available is left an open question.

Relevant to these issues is a study by Donaldson (1959). Subjects were asked to construct a set of statements which, taken together, would give sufficient information for five objects (A, B, C, D, E) to be matched one to one with five other objects (a, b, c, d, e). A positive statement such as 'A goes with a' ($A = a$, for short) scored 5 points. A negative statement such as 'A does not go with b' ($A \neq b$) scored only 1 point. Subjects had to try to make as low a score as possible, which means that negative statements had 'privileged status'. The most economical solution consisted of $[(n-1)+(n-2)+(n-3)+\ldots+1]$ negative statements which, for $n = 5$, works out at 10 negatives. For example:

(1) $A \neq b$	(5) $B \neq c$	(8) $C \neq d$	(10) $D \neq e$
(2) $A \neq c$	(6) $B \neq d$	(9) $C \neq e$	
(3) $A \neq d$	(7) $B \neq e$		
(4) $A \neq e$			

An optimal solution implies the derivation of positive statements from given sets of negative ones. Thus statements (1)–(4) lead to the derived positive statement $A = a$. From this derived positive statement can be further derived the negative statement $B \neq a$. This, in conjunction with statements (5)–(7), leads to a second derived positive statement $B = b$, etc. This sort of transformation sequence occasioned much difficulty for school children and college students alike. Some sort of 'block' typically occurred after the derivation of the first positive statement ($A = a$). Donaldson suggests that perhaps so strong a feeling of finality attaches to positive statements that subjects fail to see them as means to further ends within the framework of the problem. Positive matchings are after all the ultimate end of the endeavour. A second possibility is that negative information is, in a somewhat irrational way, distrusted as a foundation for transactions that are at all far-reaching.

With solutions of less than optimum economy, containing one or more positive statements, there is scope for transformations in the direction opposite to that considered above, i.e. from a given set of positive statements may be derived negative statements which, in turn, lead to further positives. Take the solution:

(1) $A = a$ (2) $B = b$ (3) $C \neq d$ (4) $C \neq e$ (5) $D \neq e$

From the positive statements (1) and (2) may be derived the negative statements $C \neq a$ and $C \neq b$. These, in conjunction with statements (3) and (4), enable the derivation of the positive statement $C = c$. This sort of transformation sequence occasioned no undue difficulty. The suggestion of Bruner *et al.* that avoidance of indirect procedures stems from awareness of the risk of error in the course of transforming data is accordingly rejected. There is risk of error in all transformations, yet subjects displayed ineptitude only with transformations involving the development of positive statements derived, at an earlier stage, from given sets of negative statements. Only three subjects attempted an all-negative solution straight off. All three encountered difficulty and introduced positives despite their high 'cost'. Donaldson takes this as unambiguous evidence of inability (as opposed to unwillingness) to take full advantage of the transformation possibilities of negative information and hence competently to handle indirect procedures. But fourteen of the nineteen school children and forty-nine of the ninety college students did win through to an all-negative solution in the end, which renders an unqualified 'inability' interpretation untenable. Rather would it seem that subjects are unable competently to handle indirect procedures straight off, but many develop the ability to do so with a little practice.

EXPERIMENTS WITH COMPOUND CODE ITEMS

The experiments reported here follow from those of Bruner *et al.* and Donaldson. They seek to determine (*a*) whether difficulty with and avoidance of indirect procedures extend to the solving of compound code items. If so, (*b*) what is the locus of the difficulty? And (*c*) is it inability to cope therewith, or unwillingness to make the attempt, that leads to the avoidance?

Test materials

The code items consist of five 5-letter words listed across the page, with coded versions listed vertically, in a different order, underneath. The code is arbitrary

(i.e. not based on any rule of alphabetical sequence) and solution is dependent on structural rather than semantic features of the words. The coded words are incomplete, but contain sufficient letters for the attainment of a unique and unequivocal solution. For example:

(1) FOLKS (2) TABLE (3) DRIFT (4) LIBEL (5) STEAK

–	–	–	*C*	–	( )
–	–	–	–	–	( )
Q	–	–	–	–	( )
C	–	*Q*	–	*U*	( )
U	–	–	–	–	( )

Item (i)

C occurs in first and fourth positions, and could stand for *F* (the first coded word representing DRIFT and the fourth FOLKS) or for *L* (the first coded word representing TABLE and the fourth LIBEL). *Q* occupies first and third positions, and can only stand for *L*. Hence the third coded word is LIBEL and the fourth FOLKS (and the first therefore DRIFT). Since $C - Q - U$ is FOLKS, *U* stands for *S*. The fifth coded word has a *U* in first position and so must represent a word beginning with *S*. Of the two words as yet unidentified only STEAK meets this requirement. The fifth coded word is therefore STEAK and the coded word for which no letters are given must, by default, be TABLE.

Ten items similar to Item (i) constitute Test A. Ten further items having the same sets of words but with a slightly different choice and distribution of coded letters constitute Test B. Item (i) for example becomes:

(1) FOLKS (2) TABLE (3) DRIFT (4) LIBEL (5) STEAK

–	–	–	*C*	–	( )
–	–	*X*	–	–	( )
Q	–	–	*J*	–	( )
C	–	*Q*	–	–	( )
–	–	–	–	–	( )

Item (ii)

A single *X* and a single *J* replace the pair of *U*'s from Item (i). FOLKS, LIBEL and DRIFT are identified exactly as before, but STEAK AND TABLE can no longer be identified by matching up pairs of coded letters against pairs of letters occurring in the same ordinal positions in the uncoded words since no appropriate letter-pairs are given. Turning to the single letters, since $Q - - J -$ is LIBEL therefore *J* stands for *E*. What *X* stands for is not known. If the coded counterparts of STEAK and TABLE are to be identified by other than arbitrary allocation this can, in terms of the letters given, only be accomplished indirectly. *X*, as well as being a positive instance of a class of letters unknown, is also a negative instance of the class of *E*'s. From the positive statement '*J* stands for *E*' ($J = E$) can be derived the negative statement '*X* does not stand for *E*' ($X \neq E$). The latter may alternatively be expressed in positive form as '*X* stands for *not-E*' ($X = \bar{E}$). The fact that the second

coded word has an X in third position can now be interpreted to mean that it represents a word that does not have an E in third position. This eliminates STEAK. Hence the second coded word is TABLE and the fifth STEAK, rather than vice versa. This final stage of the solution sequence is indirect in the sense in which Bruner *et al.* use the word: it involves elimination of all (here only one) alternative possibilities by showing them *not* to be the case. In addition, the elimination is achieved via transformations of data akin to those discussed by Donaldson.

The first five items of Test B are similar to Item (ii) above. Each of the remaining five items has only *one* letter-pair among the coded letters. In illustration, Item (ii) could be restructured as follows:

(1) FOLKS (2) TABLE (3) DRIFT (4) LIBEL (5) STEAK

– – – C J ( )
– – X – – ( )
Q – – – – ( )
C – – Z – ( )
U – M – – ( )
( )

Item (iii)

Solution of Item (iii) is far less direct than for Items (i) and (ii). There is a greater demand for elimination of alternative possibilities, and this entails much manipulation and transformation of data. That no item in Test B is soluble without some use of indirect procedures is in contrast with the tasks used by Bruner *et al.* and Donaldson. In their tasks direct procedures could always be used, albeit at high cost or the expenditure of excessive effort. Success via the use only of direct procedures is precluded for Test B items; the only alternative to the use of indirect procedures is arbitrary allocation.

Experiment 1

Procedure and results

Seventy-nine first-year psychology students at the University of Edinburgh were assigned at random to two groups. One group ($N = 39$) did Test A; the other group ($N = 40$) Test B. A 20 min. time limit was imposed. Both tests were prefaced by two sample items, neither of which demands any indirect procedure. (If sample items involving indirect procedures had been given straight off, and Test B not found unduly difficult, it would remain a matter of speculation whether such difficulty might have been found had sample items involving indirect procedures not been given, i.e. had subjects not been acquainted with the appropriate solution procedures and instructed in their use.) Results appear in Table 1. The difference between groups is significant at beyond the 1% level for number of whole items correct. (All P values cited come from Mann-Whitney U-tests.) The other differences, although they fail to reach the 5% level of significance, are all in the direction indicating inferior performance on Test B as regards both speed (number of words attempted) and accuracy (error ratios). That Test B, whose items call for the utilization of indirect procedures, is found more difficult than Test A (in which no such demand is made) indicates that many subjects are not able *spontaneously* to apply indirect procedures to the solving of compound code items. Experiment 2 seeks to determine whether

this is due to some difficulty inherent in the indirect procedures themselves, or merely because the subjects lacked acquaintance with and knowledge of the appropriate procedures.

Table 1. *Means, ranges and error ratios for Tests A and B in Expt.* 1

	Test A (N = 39)		Test B (N = 40)	
	Mean	Range	Mean	Range
Single words correct (maximum 50)	18·3	0–50	12·9	0–32
Whole items correct (maximum 10)	3·4	0–10	1·9	0–6
Words attempted	19·6	0–50	15·9	0–35
Error ratio (% of words incorrect)	52/764 = 6·8 %		120/636 = 18·9 %	

Table 2. *Means, ranges and error ratios for Tests A and B in Expt.* 2

	Test A (N = 19)		Test B (N = 20)	
	Mean	Range	Mean	Range
Single words correct (maximum 50)	19·6	0–50	12·8	0–30
Whole items correct (maximum 10)	3·8	0–8	2·1	0–6
Words attempted	21·6	10–42	16·4	5–32
Error ratio	39/411 = 9·5 %		73/328 = 22·3 %	

Experiment 2

Thirty-nine first-year students were assigned at random to two groups: nineteen did Test A; twenty Test B. Test B was this time prefaced by an additional sample item whose solution is partially indirect. The solution of this item was gone through *in detail*, and every effort was made to ensure that the data transformations involved were fully understood by the subjects doing Test B. Conditions were otherwise exactly as for Expt. 1. Results appear in Table 2. Performance on Test B is again inferior to that on Test A, P values being significant beyond the 5 % level both for number of words and number of whole items correct. The difference in difficulty between Tests A and B is just as pronounced as in Expt. 1, if not more so. Certainly it has not been reduced by the provision of instruction in the use of indirect procedures. (Compare the finding of Bruner *et al.* that despite detailed explanation of the nature of disjunctive concepts subjects persist in making choices and using information in ways specifically unsuited for disjunction.) It could be objected that a single sample item involving indirect procedures is scarcely sufficient to acquaint a subject therewith but, even allowing this objection, there would still seem to be difficulty associated with or inherent in the indirect procedures themselves.

Discussion

Experiments 1 and 2 have shown that lack of facility in the use of indirect procedures does extend to the solving of compound code items. And since on Test B one must either use indirect procedures or fail (apart from chance successes), the high

failure rate is strongly suggestive of inability on the part of the subjects who failed. The possibility does, however, remain that failure was preferred to the labour that indirect procedures entail. As Bartlett points out (1958, p. 84), subjects sometimes convey the impression that they cannot do anything with a problem when all that this really means is that they are not sufficiently interested to go to the trouble of working out what to do with it.

That the difficulty of the Test B items *is due* to their involving indirect procedures can be seen by comparing the error patterns for the two tests. Errors on Test A are few in number, and so distributed that no one word in any item is wrongly identified more frequently than is any other. Errors are due mainly to 'carelessness'. Subjects rush through the items and do not bother to check their solutions. Errors on Test B are far more prevalent, and no longer mainly due to carelessness. They are concentrated on words whose identification entails the use of indirect procedures, and it would seem that subjects are brought to a halt by these words and resort is had to guessing. On the first five items of Test B, where only the final pair of words cannot be identified by direct matching of coded with uncoded letter-pairs, there is a 50% chance of guessing incorrectly. This contributes to the error scores. But there is also a 50% chance of guessing correctly, which means that the number of subjects failing to work eliminatively and resorting to guessing is almost certainly greater than the error scores alone suggest. On the five remaining items of Test B the number of words not able to be decoded by direct matching procedures has increased from two to three. Since there are six possible matchings of three coded with three uncoded words, the probability of chance success is now only 1 in 6. Unfortunately the error patterns for these items could not be studied because few subjects had time to attempt them in the 20 min. allowed.

Turning now to the nature of the psychological difficulty associated with indirect procedures, an important contributory factor is a lack of appreciation that any instance is, simultaneously, a member of more than one class. All A's for example are, at the same time, $\bar{B}$'s and $\bar{C}$'s and ... $\bar{Z}$'s. Difficulty in viewing the same thing as simultaneously of two different classes has been demonstrated repeatedly by Piaget, e.g. the difficulties experienced by young children in grasping number as being both cardinal and ordinal, or in regarding daffodils as belonging simultaneously to the subordinate class 'daffodils' and the supraordinate class 'flowers'. It is not only children who experience this sort of difficulty, judging by the failure of so many first-year university students to appreciate that a letter of the alphabet—A for instance—belongs not only to the class of A's but also to the various supraordinate 'equivalence' classes $\bar{B}$'s, $\bar{C}$'s, ..., $\bar{Z}$'s that are the complementaries of the classes of the letters of the alphabet other than A itself. Why does there exist this lack of competence in dealing with classes such as the class of $\bar{X}$'s? Heidbreder (1945, 1948) argues that people prefer to work with cues that are directly perceptually apprehensible. The various $\bar{X}$'s tend not to bear any particular perceptual similarity one to the other, and this may lead to a reluctance to work with them. But the problem is less a matter of reluctance than one of failure even to conceive of such a class as that of $\bar{X}$'s, a class which, when united with the X's, exhausts the contents of the supraordinate class 'letters of the alphabet'. Linguistic factors may have something to do with this. Classes such as $\bar{X}$'s are purely verbal, the only feature the various

class members have in common being that they can all be labelled '$\bar{X}$'. In at least one experiment (Bruner & Olver, 1963) linguistic convention has been found to be used but infrequently as a basis of classification.

A second major factor contributing to the difficulty of Test B items is their demand for transformations involving a change of logical subject. Take Item (ii) above. The transition from $J = E$ to $X \neq E$ involves the derivation, from a statement about J, of a statement about something other than J. The opposite transition from $E = J$ to $E \neq X$ is probably easier. Both are statements about E. But this transition, while equivalent to the former from the point of view of formal logic, is not the one that subjects doing the code items actually attempt. Donaldson (1956) tentatively suggested that transformations of data are the more difficult if they involve a change of logical subject. The present results substantiate this. Donaldson's further suggestion (1959) that a feeling of finality may attach to positive statements is also relevant here. Such a feeling would contribute to failure with the seemingly simple transition from $J = E$ to $X \neq E$, the positive $J = E$ being regarded purely as an end in itself rather than as also a means to a further end within the framework of the problem.

Conclusions

The experiments with compound code items reported in this paper provide further support for the suggestion of Bruner *et al.* (1956) that lack of facility with indirect procedures may extend to a variety of cognitive activities other than categorization. Many subjects seem truly unable—rather than merely unwilling—to employ the indirect procedures appropriate to the solving of the Test B items. It is suggested that the difficulty experienced with these items stems from the need to work with equivalence classes that are negatively defined and purely verbal, and/or the demand for data transformations involving a change of logical subject.

References

Bartlett, F. C. (1958). *Thinking. An Experimental and Social Study.* London: Unwin.

Bruner, J. S., Goodnow, J. J. & Austin, G. A. (1956). *A Study of Thinking.* New York: Wiley.

Bruner, J. S. & Olver, R. R. (1963). Development of equivalence transformations in children. *Monogr. Soc. Res. Child Develpm.* **28**, 125–41.

Donaldson, M. C. (1956). The relevance to the theory of intelligence testing of the study of errors in thinking. Ph.D. thesis, Edinburgh University Library.

Donaldson, M. C. (1959). Positive and negative information in matching problems. *Brit. J. Psychol.* **50**, 253–62.

Heidbreder, E. H. (1945). Towards a dynamic psychology of cognition. *Psychol. Rev.* **52**, 1–22.

Heidbreder, E. H. (1948). The attainment of concepts: VI. Exploratory experiments on conceptualisation at perceptual levels. *J. Psychol.* **26**, 193–216.

(*Manuscript received* 13 *January* 1964)

Brit. J. Psychol. (1965), 56, 1, *pp.* 53–59
Printed in Great Britain

PATTERNS OF REACTION TIME RESPONSES

BY MYFANWY KERR*, ROSEMARY MINGAY†
AND ALICK ELITHORN*

Institute of Neurology and the Royal Free Hospital, London, W.C. 1

The reaction time responses of forty-four subjects to paired stimuli separated by an interval of 100 msec. were analysed to test whether there was evidence for several distinct types of response patterns, and, if so, whether these patterns were associated with individual performances.

The analysis confirmed both these hypotheses, six main types of response being distinguishable and the response patterns being markedly different for different individuals.

In a previous paper the patterns of reaction times to pairs of stimuli separated by a short interval of 100 msec. were shown to be related to the subject's laterality preferences (Kerr, Mingay & Elithorn, 1963). The responses of ten right-handed and ten left-handed subjects were analysed. It was shown that responses in which both reaction times (RT 1 and RT 2) were undelayed occurred more frequently when the first stimulus occurred on the dominant side.

When only RT 1 was undelayed the magnitude of the delay in RT 2 was significantly greater when the second stimulus was on the dominant side. Other types of response showed no obvious relationship with hand dominance.

It was clear from inspection of the data, however, that the different types of response were not randomly distributed amongst individual subjects, but that most subjects showed a marked preference for one type of response. The scope of the present paper is to classify in more detail the responses made by the whole group of forty-four subjects and to see to what extent responses by individual subjects tend to fall into one classification.

APPARATUS AND EXPERIMENTAL DESIGN

Details of the apparatus and experimental technique were described in the previous paper. Briefly, subjects were required to make key pressing responses to visual stimuli, occurring on the left or right, singly or in pairs, the latter being either simultaneous or separated by a time interval of 100 msec. As a control, alternate runs contained single stimuli only, the subject being told that there would be only one stimulus in each situation, but not the side on which it would occur. The first two runs, one control and one experimental, were treated as practice and excluded from subsequent analysis. Each subject was instructed to respond to every stimulus as quickly as possible.

The subjects were forty-four healthy adults, eighteen females and twenty-six males, unpractised in this type of experiment, who were drawn mainly from the medical and other staff of the National Hospitals for Nervous Diseases, London, and from among the medical students of the Royal Free Hospital Medical School, London. The age range was from 19 to 41 years (mean age 27·5) with a slight bias towards the younger age groups.

* Member of the Medical Research Council external staff.
† Former member of the M.R.C. Neurological Research Unit, National Hospital for Nervous Diseases, London.

RESULTS

The responses to paired stimuli separated by a 100 msec. time interval, which are the subjects of this analysis, were surprisingly varied. The simplicity of the task is deceptive. Where two stimuli are closely ordered in time it is extremely difficult to respond without delay to the first and avoid an apparently intractable delay in responding to the second (Telford, 1931; Craik, 1947, 1948; Vince, 1948). Neglecting the possibility that one stimulus may facilitate the other,* if subjects are obeying the instruction to respond to each stimulus as quickly as possible, two response patterns are possible. A first reaction time which is of 'normal' duration is accompanied by either a normal or 'longer than normal' second reaction time; in other words the second response is either (i) independent of the first or (ii) delayed in relation to the first, the latter being by far the more commonly observed response. The responses in the present experiment, however, were more often of a 'grouped' type in which the two stimuli were treated as a unit. Grouping can take several forms, but a factor common to all is that no response is initiated to either stimulus until both stimuli have been perceived; thus (i) the subject can depress both keys together, in which case RT 1 exceeds RT 2 by 100 msec., the length of the inter-stimulus interval (the grouped response as defined by Welford, 1959); (ii) the keys can be depressed in the right order and with the inter-stimulus interval simulated more or less successfully so that RT 1 and RT 2 are roughly equal but both are longer than normal ('sophisticated grouping'); (iii) the keys can be depressed in the right order but with RT 2 even longer than RT 1 ('long' delayed responses); (iv) finally the keys can be depressed in the wrong order with RT 2 quicker than RT 1 by more than 100 msec. (reversed responses). These six categories account for 93% of the observed responses. They are shown in Table 1 where they are separated into Classes I and II on the basis of the criteria that are given in detail below.

Table 1. *Classification of response patterns*

Class	Type	RT 1	RT 2	No. of responses	% of total
I Non-grouping response patterns	(1) Independent	Normal	Normal	83	11
	(2) Delayed	Normal	> Normal	226	29
II Grouping response patterns	(3) Grouped	> Normal	≃ RT 1 − 100 msec	76	10
	(4) Sophisticated grouping	> Normal	≃ RT 1	81	10
	(5) 'Long' delayed	> Normal	> RT 1	222	28
	(6) Reversed	> Normal	< RT 1 − 100 msec.	44	5

It is clear that, for adequate definition, response patterns must be classified in terms of two relations. The first relation is that between RT 1 and RT 2 separately and some criterion of 'normality' for a single response for the appropriate time and side of stimulation. The second relation is that between the reaction times themselves (RT 1 and RT 2). As regards the first criterion, in this type of experiment where there is considerable inter-subject variability, normality is only meaningful in relation to each individual's performance and the most reliable control is obviously

* In the present experiments facilitatory effects are not marked and cannot be demonstrated unequivocally. This does not affect the thesis put forward in this paper.

that given by response times to single stimuli. It seems preferable also to base the criterion on single response times from the 'control' runs rather than from the experimental runs, since in the latter the admixture of paired responses will affect subjective expectancy, and hence response times, even in the single stimulus situation. The proposed control is not quite perfect because of changes in stimulus probability between the two stimulus times, but it is unlikely that this effect is of any practical importance. The control is probably the best that can be devised within the experimental framework.

It was therefore decided to regard as normal for each subject those times which were within his control range of single reaction times, excluding the longest. Exclusion of the longest time is a safeguard against the occasional atypical long response. As a safeguard against 'concealed' premature responses an absolute value of 175 msec. was taken as the lower limit for all subjects. It is worth mentioning that the observed number of premature responses (i.e. those where a response is made when there is no stimulus) is very low (1 % of all single responses in either the control or the experimental situation).

The second criterion is needed to define more exactly the different types of grouping which may occur. Where a response is defined in terms of a specific relation between RT 1 and RT 2 (e.g. grouped responses where RT 2 = RT 1 − 100 msec.) an error of ± 30 msec. has been allowed. Thus a response where RT 1 = 300 msec. and RT 2 = 230 msec. (for a subject whose normal response range was below 300 msec.) would be classified as grouped.

Examination of the data classified according to these criteria showed a distinct tendency for different subjects to favour different types of response. In some cases this was so marked that one response pattern would be favoured almost to the exclusion of any other.

Figs. 1*a* and *b* are scattergrams of the reaction times for five subjects in the left/right and right/left order of stimulus presentation, respectively. The subjects have been chosen because the patterns of preferred responses contrast markedly. All the responses for each subject have been plotted and a line drawn to enclose the most typical in each case.

The tendency for a subject to favour a certain type of response is clearly shown. In the left/right situation (Fig. 1*a*) subject 1 shows a majority of delayed responses; subject 2 shows a tendency towards independent responses near the line of equality where RT 1 = RT 2; subject 8 shows mainly grouped responses with RT 1 approximately 100 msec. longer than RT 2; subject 30 shows sophisticated grouping, and subject 44 long responses with RT 2 even longer than RT 1.

In the right/left order of stimulus presentation (Fig. 1*b*) subjects 2, 8 and 30 show virtually the same response patterns as in the left/right situation. However, the response patterns for the other two subjects, both of whom are right handed, show a dominance effect. Subject 1 makes responses much nearer the line of equality in the right/left situation than in the alternative one. This tendency to make more independent responses in the dominant/non-dominant situation reached a satisfactory level of significance in the group analysis and has been reported in a previous paper (Kerr *et al.* 1963). Subject 44 shows a dramatic shift from long responses with RT 2, even longer than RT 1, to reversed responses in the right/left situation. This subject

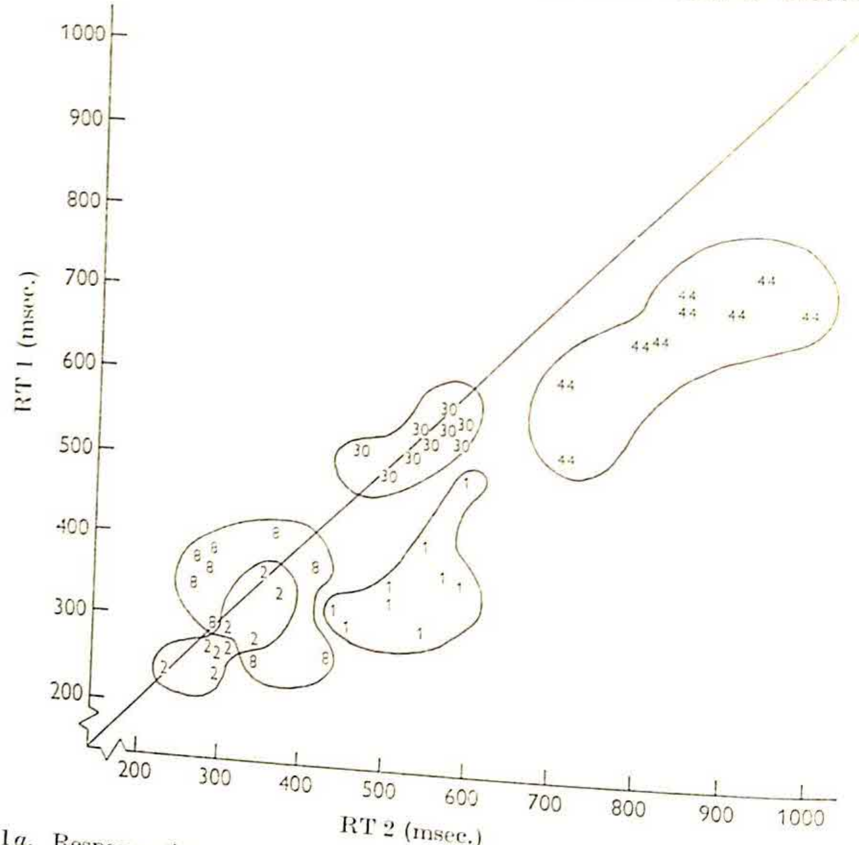


Fig. 1*a*. Response times for 5 subjects in the left right stimulus–response situation.

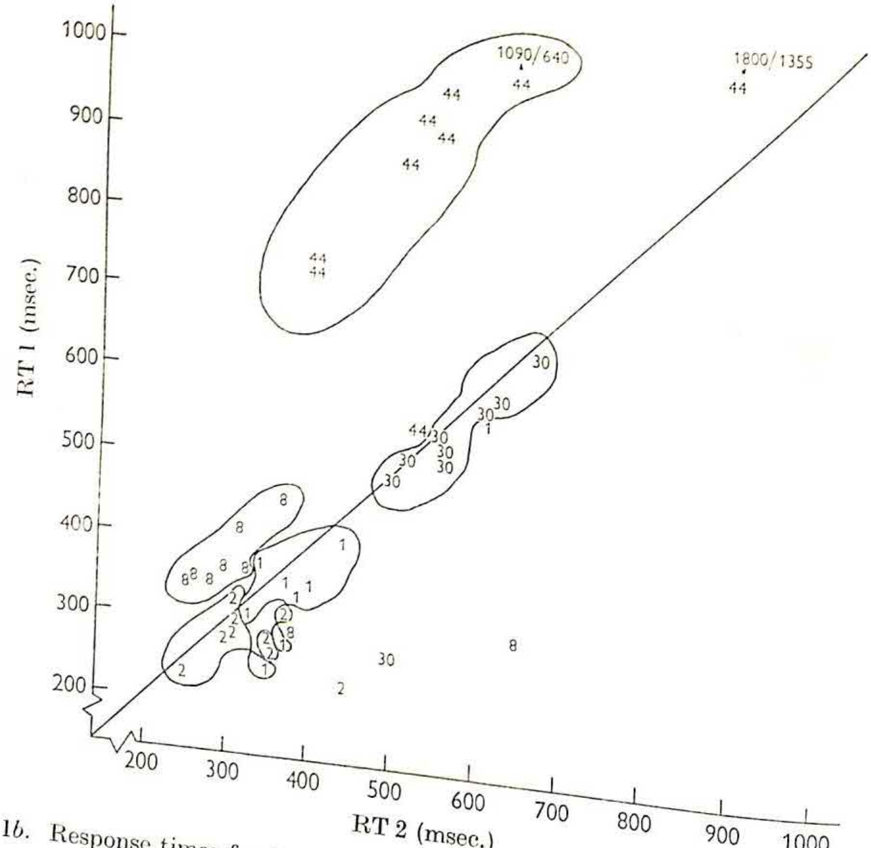


Fig. 1*b*. Response times for 5 subjects in the right left stimulus–response situation.

in fact treated all the stimuli separated by 100 msec. as though they were given in the left/right sequence.

If apparently individual response preferences as marked as this can be shown not to be randomly occurring groupings, are they sufficiently clear cut to be used as a basis for subject classification? Considering in the first place the incidence of Class I and Class II responses with the forty-four individual subjects, the value of χ^2 in the resulting 2×44 contingency table ($\chi^2 = 222$; D.F. 43, $P < 0{\cdot}001$) provides confirmation that the distribution of these two types of response is not homogenous and that the concept of individual preferences has statistical validity. Clearly our subjects can be divided into those who show a relative preference for Class II responses and those who do not. Of the forty-four subjects tested only three made responses that were all of the same type—either Class I or Class II. The data are insufficient to determine whether these subjects are merely extreme examples of two preferences, or whether the subjects should be grouped into those who mainly make responses of one or other type and a third group whose strategies are mixed.

The classification of responses into Classes I and II is open to the objection that is a very broad one. Into Class I fall not only those responses in which RT 2 is normal as well as RT 1, but also those in which RT 2 is markedly abnormal. Class II comprises not only grouped responses but also long delayed responses, sophisticated grouping and reversed responses.

Each class was therefore broken down into its subclasses and the same technique used to test for homogeneity. For Class I (twenty subjects) the frequency distribution of independent and delayed response yielded a highly significant value for χ^2 (83·12, D.F. 19, $P < 0{\cdot}001$). Evidently some subjects tend to delay their response to S 2 while others are capable of independent responses in this situation. In fact seven (35 %) of the twenty subjects accounted for fifty (83 %) out of the sixty independent responses made by this group. These seven subjects form a group whose ratio of independent to delayed responses is higher than unity (5:4) and much higher than the mean ratio (1:3) for the group as a whole.

The analysis of the Class II responses is more complicated. When the data are cast into four subclasses (grouped, sophisticated grouping, long delayed, or reversed) the resultant table contains many cells of low or zero frequency. The significance of the value of χ^2 (338·4, D.F. 63) was therefore assessed by evaluating the exact mean and variance of χ^2 and referring the critical ratio $\{\chi^2 - E(\chi^2)\}/V(\chi^2)$ to the normal curve (Bartlett, 1937; Dawson, 1954). The level of significance was $P < 0{\cdot}001$. Evidently in this class also performances are not homogeneous. A satisfactory classification of subjects according to a preferred type of response pattern, however, is difficult to achieve in this experiment; with a small sample of responses from each subject and the existence of four possible classes of response, the frequency of a given type of response would have to depart markedly from the mean expectation to provide any valid classification. A better classification of individuals could be obtained by an iterative technique involving the calculation of expected frequencies for each subgroup which would minimize the value of χ^2, and thus demonstrate maximum homogeneity within each subgroup. It seems unwise, however, to overload the present data with an elaborate analysis which should await a larger sample. For the moment it is sufficient to demonstrate the practicability of such an analysis.

Discussion

Earlier work on reaction times to visual stimuli has suggested that the responses to paired stimuli closely following each other were limited to delayed responses and what were broadly termed 'grouped' responses. No evidence of independent response times was found and grouped responses were analysed no more closely than to say that the motor responses in most of the grouped pairs of reactions were not simultaneous but successive, although often 'M_2 could begin before M_1 had finished' (Welford, 1959). The present study suggests that six distinct types of response can be defined with some degree of precision at the time interval studied, and that the occurrence of a particular response is not random but reflects a pattern of behaviour on the part of the subject. This behaviour pattern is largely within the voluntary control of the subject, the main exception being the non-grouped responses where the preference for a delayed type of response rather than an independent one appears to be determined by some physiological limitation peculiar to the individual. Grouped responses on the other hand appear to represent a strategy chosen by the subject to 'make the best of a bad job', that is to say his inability to respond to each stimulus as soon as it appears. The present study provides no evidence of any correlation between these types of response and personality traits. A further study to investigate this point is now being undertaken.

One point of interest peripheral to the main object of the experiment has emerged from the data. In relation to previous related experiments Welford (1959) maintained that, in the absence of grouping, a delay in the second response to two stimuli closely ordered in time was unavoidable. The results of experiments by Elithorn & Lawrence (1955), and Halliday, Kerr & Elithorn (1960) do not confirm this finding. Their data suggested that, with practice at any rate, delay in RT 2 was not invariable, the implication being that the response becomes more of a reflex action with increasing practice. This concept links up with the work of Leonard (1958), who has shown that in a situation with high stimulus–response compatibility independence is the rule rather than the exception. The present experiment confirms that even with unpractised subjects independence can occur with a frequency which precludes its explanation in terms of premature responses. Furthermore, the incidence of independent responses varies markedly with individual subjects.

References

BARTLETT, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. Roy. Soc.* A, **160**, 268–82.

CRAIK, K. J. W. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.* **38**, 56–61.

CRAIK, K. J. W. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.* **38**, 142–8.

DAWSON, R. B. (1954). A simplified expression for the variance of the χ^2 function on a contingency table. *Biometrika*, **41**, 280.

ELITHORN, A. & LAWRENCE, C. (1955). Central inhibition—some refractory observations. *Quart. J. Exp. Psychol.* **7**, 116–27.

HALLIDAY, A. M., KERR, M. & ELITHORN, A. (1960). Grouping of stimuli and apparent exceptions to the psychological refractory period. *Quart. J. Exp. Psychol.* **12**, 72–89.

LEONARD, J. A. (1958). Partial advance information in a choice reaction task. *Brit. J. Psychol.* **49**, 89–96.

Kerr, M., Mingay, R. & Elithorn, A. (1963). Cerebral dominance in reaction time responses. *Brit. J. Psychol.* **54**, 325–36.
Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *J. Exp. Psychol.* **14**, 1–36.
Welford, A. T. (1959). Evidence of a single channel decision mechanism limiting performance in a serial reaction task. *Quart. J. Exp. Psychol.* **11**, 193–210.
Vince, M. A. (1948). The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.* **38**, 149–57.

(*Manuscript received 7 March* 1964)

The following experiment is mainly intended to provide such information on two points. First, does the interaction between information and compatibility already noted still appear in a fresh situation? Secondly, if reaction time is being affected by other variables, notably a second distracting task or uncertainty about the time of arrival of stimuli, do these effects also interact with compatibility?

With regard to the first purpose, information can of course be varied either by changing the number of alternative stimuli, or by varying the relative probability of different stimuli in a fixed vocabulary. The first of these has been studied by Brainard *et al.* (1962), and the second by Fitts *et al.* (1963). In both cases, however, the stimuli were visual and the responses either vocal or manual, and compatibility was varied by changing the whole set of responses used for a particular set of stimuli. In the present experiments the stimuli were tactual and the responses manual. Compatibility was varied by keeping the same set of responses but changing the particular response appropriate to a particular stimulus.

As incidental purposes, the experiment was so designed that it would shed light upon the interaction between number of alternative stimuli and uncertainty about time of arrival of stimuli. It was also arranged that performance on a stimulus occurring with a particular fixed probability could be compared when it was embedded in different sequences made up of two different numbers of other alternative stimuli.

Method

The skeleton of the experiment was that each subject worked for two sessions on separate days; he performed a 2-choice task on one day and a 4-choice task on the other. In other respects conditions were constant for each man, but different subjects performed compatible and incompatible tasks under each of four different conditions: (i) with prepared reactions; (ii) with unprepared reactions; (iii) with prepared reactions and distraction; and (iv) with prepared reactions and biased distribution of signals.

Apparatus

This was a much modified version of that described by Leonard (1959). The subject rested the tips of his fingers on specially constructed keys, the position of which could be adjusted to suit his hands. Each key consisted of a ring mounted on a pivoted arm with a spring return and a stop below the ring. Through the centre of the ring projected the vibrator which delivered the signal. The vibrator consisted of a 'Tufnol' rod of 6 mm. diameter, attached to an armature which was energized at 50 cyc./sec. by a surrounding coil mounted between the poles of a permanent magnet. The whole vibrator assembly was pivoted and the pressure of the rod on the subject's finger was controlled by a counterbalancing weight at a value of approximately 20 g. The pressure required to operate the response keys was 80–90 g. The experimenter selected the appropriate vibrator by means of one of four switches and then initiated stimulation by the depression of a single common switch. This operation also started a 'Dekatron' timer which timed the reaction to the nearest 10 msec. Depression of any one of the keys operated two microswitches which arrested the vibration and the 'Dekatron'. When a wrong key was depressed a single lamp was lighted on the experimenter's panel indicating an error but not establishing the nature of the error.

Procedure

(*a*) *Factors common to all four major conditions.* The two sessions for each subject were on successive days. On one day each subject made 2-choice reactions using either both index fingers or both middle fingers, while on the other day he made 4-choice reactions using both the

index and the middle fingers from each hand. In the 2-choice reactions, half the subjects responded with both index fingers and the other half with both middle fingers. Subjects who were to respond in a *compatible* fashion were instructed to press down the finger that was stimulated by the vibration. The remaining subjects, who were to react in an *incompatible* fashion, were told to press down the corresponding finger of the other hand, i.e., if the index finger of the left hand was stimulated, the subject had to respond by pressing down with the index finger of the right hand. Within each group half the subjects had 2-choice reactions on the first day and the other half had 4-choice.

In each session there were 64 trials. In the 4-choice session the four stimuli occurred equally often in each sixteen trials, and no stimulus occurred on more than two consecutive trials; in the 2-choice session the stimuli occurred equally often in each group of eight trials, and runs of three but no more were present. (These restrictions do not apply to Major Condition 4, that of bias.)

Before each session the subject was instructed to respond as fast as he could without making errors, and that if he did make an error he would be told that he had done so immediately afterwards. He was given a practice trial on each of the stimuli he was going to have in the succeeding session, and, if he made an error in response, that signal was repeated. A total of a hundred and seven subjects were tested, all of whom were naval ratings between the ages of 18 and 30. They were divided as shown in Table 1 between the various conditions.

Table 1. *Mean response times (msec.)*

Condition	Incompatible N	Incompatible 2-choice	Incompatible 4-choice	Incompatible Difference	Compatible N	Compatible 2-choice	Compatible 4-choice	Compatible Difference
Prepared	12	337	530	193	10	261	351	90
Unprepared	12	500	643	143	11	318	395	77
Distraction	10	542	662	138	12	342	356	14
Bias:								
Biased responses	8	343	410	67	12	248	300	52
Other responses	8	351	559	208	12	242	337	75

(*b*) *Differences between major conditions.* (1) Prepared reaction. The interval between the stimuli was 10 sec. and the subject was given an oral warning about 2 sec. before a signal was presented.

(2) Unprepared reaction. The reaction signals occurred without warning at intervals of 10, 20, 30 or 40 sec. after the last reaction signal, each interval occurring equally often within each quarter of the 64 trials.

(3) Prepared reaction with distraction. Again the interval between the signals was 10 sec., but in this case the warning signal which preceded the reaction signal was a spoken letter of the alphabet. These letters were spoken on tape, a different one every 5 sec.; so each alternate letter acted as a warning signal. Within a group of ten letters or five reactions, one letter was repeated; at the conclusion of the group, the subject was asked to say aloud which letter had occurred twice. The letters used were the first ten letters of the alphabet, excluding D. The tape was played on a 'Ferrograph' recorder.

(4) Prepared reaction with biased distribution of signals. The procedure was the same as in (1), except that one signal occurred on three-quarters of the occasions: in the 2-choice series one stimulus occurred on 48 trials while the other occurred on only 16 trials; in the 4-choice series one stimulus again occurred 48 times, but each of the other three on only five occasions. Each subject had the bias favouring the same signal for both sessions, and was informed at the beginning of the first session that that signal would be more frequent. Over the group as a whole the bias was equally distributed between the four possible fingers.

Scoring

The scores were taken only from the second half of each session. In conditions 1, 2 and 3 the score taken in the 4-choice reaction was the mean of the medians of the times of the last five correct responses made with each of the fingers corresponding to those used by the subject in his 2-choice session, i.e. either both index fingers or both middle fingers. In the 2-choice reaction, the score was the mean of the medians of the times of those five responses with each finger which had the same positions in the series as those used in scoring the 4-choice reaction.

In condition 4, that of bias, in both the 2-choice and 4-choice situations, the score taken for the response time to the frequent signal was the median of the times of the last nine correct responses to that signal. For the infrequent signal or signals, the score taken was the median of the times for the last seven correct responses; if seven were not available, in the second half of the session, then the median of the last five was taken. In the 4-choice this score included response times to all three infrequent signals.

Unless otherwise stated all statistical tests are two tailed; differences between groups of subjects have been tested by Mann–Whitney U test and differences within subjects by a sign test on the number of subjects showing the effect. Two interactions, indicated in the text, were tested by analysis of variance.

RESULTS

The mean of the response times (as defined above for each subject) for each condition is shown in Table 1. The findings from the analysis of response times may be described under two headings: interactions of compatibility with other variables, and differences within one level of compatibility. Although errors are regarded as of prime importance in decision theory, their incidence in this experiment was too slight to allow analysis.

Interactions of compatibility with other variables

(*a*) With the number of alternative choices. In all conditions except that of bias (which is of course by nature anomalous), the difference between the response times in the 4-choice and 2-choice situations is significantly greater for the group making incompatible responses than for the group making compatible responses ($P < 0.05$ in each case).

(*b*) With time uncertainty. The lengthening of response time brought about by uncertainty about the time of arrival of the stimuli is greater when the subjects are making incompatible responses, at least in absolute terms. It may not be more than proportional to the longer reaction time met in the incompatible case.*

(*c*) With distraction. There is a greater effect of distraction upon response time when subjects are responding in an incompatible fashion. In this case the effect is even larger than proportional to the longer reaction time, $P < 0.05$ for each degree of choice by analysis of variance after logarithmic transformation. In addition, the mean score for performance on the distraction tasks are shown in Table 2; it will be seen that performance was better in subjects making compatible responses ($P < 0.02$).

Table 2. *Percentages of letters correct in the distraction task*

	Incompatible	Compatible
2-choice	59·1	84·9
4-choice	49·1	72·8

(*d*) With bias. When considering response times to the signal receiving favourable bias, there is no interaction between compatibility and the number of alternatives.

* There is a slight difficulty in testing the difference statistically, since all the conditions concerned were tested on different subjects so that a conventional U test cannot be applied to the interaction. On the other hand, the variance is significantly non-homogeneous so that analysis of variance can only be performed after a logarithmic transformation of the raw data. Testing an interaction after such a transformation is not equivalent to testing whether the difference of the two larger times is the same as the difference of the two smaller ones; rather it tests whether the differences are proportional to the means. It is not therefore too surprising that the interaction fails to be significant in the 4-choice condition ($F = 2.35$; at $P = 0.05$, $F = 4.08$), although it is comfortably so in the 2-choice condition ($F = 7.86$, $P < 0.01$).

The difference between the response times to the frequently occurring signal and those to the other signals interacts with compatibility in the 4-choice situation ($P < 0{\cdot}002$) but not in the 2-choice (where the difference in probability is of course smaller). The direction of the effect is that the effect of signal probability is larger when the subject is responding in an incompatible fashion.

Findings within one level of compatibility

(*a*) No significant interaction was found between time uncertainty and number of alternatives. Comparison of the difference in mean response times between the 4-choice and 2-choice situations in the prepared and unprepared condition shows no significant change in this difference. There is, however, a non-significant flattening of the slope of response time against number of alternatives in the unprepared case ($P > 0{\cdot}10$). One cannot unfortunately regard the results as disproving the interaction completely, although they certainly provide no definite evidence for it.

(*b*) Interaction between distraction and number of alternative responses.—Table 2 shows that, for both groups of subjects responding compatibly or incompatibly, performance on the distraction task was worse in the 4-choice situation than in the 2-choice ($P < 0{\cdot}03$ for both groups combined).

(*c*) Effect of number of alternative responses upon a response of fixed probability.—Response times to the signal which occurs on 75% of the occasions are longer when the remaining 25% of signals are of three kinds than when they are all the same. This difference between the 2-choice and 4-choice situations is significant for both groups of subjects ($P < 0{\cdot}02$ in each case).

Discussion

The interactions of compatibility

These results clearly support the central position which is given to S–R compatibility by theories which regard choice reaction as a process analogous to statistical decision. In the first place, the results confirm the indications from earlier work that reductions in the probability of a stimulus have a more serious effect when the correct reaction to that stimulus is not obvious. This is true both for an increase in the 'vocabulary' of equally probable signals, and for a bias in probability between signals of a fixed vocabulary. Moreover, it also appears to be true when the probability of a signal at each particular instant is lowered by dispensing with a warning signal.

A point of particular interest is the effect of low compatibility in creating difficulties for the combined performance of two tasks. Not only is the reaction itself more likely to be slow, but the other task is also more impaired. A body of existing knowledge shows that the limits of human performance do not depend upon the amount of stimulation reaching the senses for action at one time, but rather upon the amount of information which those stimuli represent. It has been further noted (Broadbent, 1958, p. 78) that lowering the discriminability of the stimuli themselves will reduce the number of tasks a man can perform simultaneously, even if the information in each task is left constant. It must now be added that difficulty in the choice of appropriate response in the task] has the same effect. The limit of

performance is indeed in some sense a limit to the processing of information, to the selection of one of a set of responses as appropriate to one of a set of stimuli: but the size of the sets is only one of the factors involved.

The question of time uncertainty

The problem of time uncertainty is that it may or may not have its effect through the same mechanism as other changes in the probability of signals. One of the classic papers in the field, that of Hick (1952), contends that reaction time is related not simply to the logarithm of the number of equiprobable alternatives, but rather to that number plus one. The rationale of this formulation is that the subject is not only choosing between the various possible overt reactions, but also, because of uncertainty about the time of arrival of the stimulus, considering the possibility of no reaction if no stimulus is present. A normal degree of time uncertainty corresponds to one extra alternative, and greater or less degrees to rather more or rather less than one. A more widely held view has been that of regarding reaction time as made up of two components, one being the logarithm of the number of alternatives, and the other independent of that number. The latter component might be supposed to depend upon time uncertainty as well as upon purely mechanical constant times in the reacting system. Welford (1960) in reviewing the literature presents arguments for adhering to Hick's original formulation.

If now the effect of time uncertainty is equivalent to the addition of extra alternative responses, it should diminish the difference between 2-choice and 4-choice reactions. This is because the effect of adding two further alternatives becomes less as the number of alternatives already under consideration is greater. Accordingly Hick's original formulation should predict an interaction between time uncertainty and the effect of the number of alternatives; the alternative view should predict no interaction.

It is unfortunate that the results are not more clear-cut. Although there is no significant interaction it does not fall short of significance by a large amount. It must be added, however, that another group of subjects tested under slightly different conditions as part of the preliminary experimentation showed even less effect. At present therefore this evidence is slight but not completely convincing evidence against the treatment of time uncertainty as equivalent to additional alternatives.

The effect of number of alternatives on a response of constant probability

It has of course been known since the work of Hyman (1953) that signals which are not equiprobable show reaction times which are not exactly related to the information conveyed by each individual signal, even though on average the reaction time for the set of signals follows the information conveyed by the set. The present result emphasizes this point further, since the information content of a stimulus remains constant if its probability is constant. On the other hand, the model of reaction time based on statistical decision supposes that reaction occurs when sufficient evidence in favour of one alternative has been accumulated by the decision mechanism. The amount of evidence needed, and therefore the reaction time, depend upon the prior probability of that reaction being correct; but, even for a fixed prior probability, the evidence needed will be greater if there is a larger number of other possible reactions.

Thus this result gives general support to the decision theory approach. Indeed. all these results do so, although they do not entail any specific version of that approach. Rather they display relations which a successful theory must explain.

References

Brainard, R. W., Irby, T. S., Fitts, P. M. & Alluisi, E. A. (1962). Some variables influencing the rate of gain of information. *J. Exp. Psychol.* **63**, 105–10.

Broadbent, D. E. (1958). *Perception and Communication.* London: Pergamon.

Broadbent, D. E. & Gregory, M. (1962). Donders' B- and C-reactions and S-R compatibility. *J. Exp. Psychol.* **63**, 575–8.

Conrad, R. (1962). Practice, familiarity and reading rate for words and nonsense syllables. *Quart. J. Exp. Psychol.* **14**, 71–6.

Crossman, E. R. F. W. (1956). The information capacity of the human operator in symbolic and non-symbolic control processes. In Draper, J. (ed.), *The Application of Information Theory to Human Operator Problems.* W.R. (D) Report, No. 2/56, Ministry of Supply.

Davis, R., Moray, N. & Treisman, A. (1961). Imitative responses and the rate of gain of information. *Quart. J. Exp. Psychol.* **13**, 78–89.

Fitts, P. M. & Seeger, C. M. (1953). S–R compatibility: spatial characteristics of stimulus and response codes. *J. Exp. Psychol.* **46**, 199–210.

Fitts, P. M. & Deininger, R. L. (1954). S–R compatibility: correspondence among paired elements within stimulus and response codes. *J. Exp. Psychol.* **48**, 483–92.

Fitts, P. M. & Switzer, G. (1962). Cognitive aspects of information processing: I. The familiarity of S–R sets and subsets. *J. Exp. Psychol.* **63**, 321–9.

Fitts, P. M., Peterson, J. R. & Wolpe, G. (1963). Cognitive aspects of information processing: II. Adjustments to stimulus redundancy. *J. Exp. Psychol.* **65**, 507–14.

Griew, S. (1958). Information gain in tasks involving different stimulus-response relationships. *Nature, Lond.*, **182**, 1819.

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. Exp. Psychol.* **4**, 11–26.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *J. Exp. Psychol.* **45**, 188–96.

Laming, D. R. J. (1962). A statistical test of a prediction from information theory in a card-sorting situation. *Quart. J. Exp. Psychol.* **14**, 38–48.

Leonard, J. A. (1959). Tactual choice reactions. *Quart. J. Exp. Psychol.* **11**, 76–83.

Mowbray, G. H. & Rhoades, M. U. (1959). On the reduction of choice-reaction times with practice. *Quart. J. Exp. Psychol.* **12**, 193–202.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, **25**, 251–60.

Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics*, **3**, 189–229.

(*Manuscript received* 16 *March* 1964)

Brit. J. Psychol. (1965), **56**, 1, *pp.* 69–75
Printed in Great Britain

CHANGES IN IMMEDIATE MEMORY WITH AGE

BY ANDREW McGHIE, JAMES CHAPMAN
AND J. S. LAWSON
Royal Dundee Liff Hospital

A test allowing assessment of the differential changes in auditory and visual short-term retention was given to groups of young, middle-aged, and elderly adults. The results indicated that all subjects found it more difficult to handle visually, as opposed to aurally, presented information. The decline in visual retention was found to increase with age, this being most noticeable after the age of 60 years. The theoretical and practical consequences of these findings are briefly discussed.

The study reported in this paper is derived from a series of clinical–experimental investigations of alterations in selective attention and perception occurring in mentally ill patients (McGhie & Chapman, 1961; Chapman & McGhie, 1962; Lawson, McGhie & Chapman, 1964). In a previous report (Chapman & McGhie, 1962) it was demonstrated that the short-term memory of schizophrenic patients is particularly vulnerable to interference by stimuli irrelevant to the task in hand. It was argued that the relative inability of schizophrenic patients to screen out extraneous information results in an overloading of short-term memory, causing these patients to perform poorly in tasks which demand accurate perception and recall. Further experimental work was carried out to assess to what extent this deficiency in short-term memory varied with the sensory modality through which the information was processed. In following up this line of inquiry, it was decided to include in our test battery a technique first reported by Buschke (1962) in his studies of auditory and visual interaction in the immediate memory of normal adult subjects. Buschke utilized the fact that if a subject is asked to note the repetition of any number in a set of random numbers (in the range 1–20) presented sequentially, then the number of correct responses provides an estimate of the amount of information registered for short-term retention. In this type of experimental design where the subject has only to note the occurrence of a repetition of one number in a series, the unrepeated numbers automatically act as distracting stimuli. In Buschke's experiment each set was composed of a sequence of 9 consecutive visual and 9 consecutive aural numbers. The presentation of each number pair, one visual and one aural, was synchronous. The subject's task was merely to press a button to indicate his recognition of the repetition of any number already presented in the sequence. Buschke presented his test to a group of 36 normal adult subjects who were 18–38 years old, and analysed the differences between the auditory and visual components of immediate memory. His main finding was that there was a much more pronounced decline in the temporary storage of visual, as opposed to auditory, information.

Buschke's procedure seemed to offer a useful method of investigating the differential changes in short-term memory in mental illness, and we therefore decided to use a modified version of this test in our work with psychiatric patients. Before applying our form of the test to clinical groups it was necessary to obtain normative data, and it is the findings from this control study which form the basis of the present paper.

analysis indicated that the fall-off with age was not uniform among the different conditions ($F = 2{\cdot}98$; D.F. 18, 764; $P < 0{\cdot}005$). In order to examine further this variation with age, we divided the 198 subjects into three age-groups: group A, under 40 years; group B, 41–60 years; group C, 61–80 years. The results are shown in Fig. 1; mean scores and standard deviations are given in Table 3, together with the results of comparing groups A and B, and B and C.

Table 3. *Mean score and standard deviations under each of the four conditions AA, AV, VV and VA for the three age-groups A (18–40 years, n = 121), B (41–60 years, n = 40) and C (61–80 years, n = 37); comparison of differences between means for groups A and B, and groups B and C*

Condition	Group A		Group B		Group C		Group A *vs.* B		Group B *vs.* C	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	*t*	*P*	*t*	*P*
AA	8·49	0·62	8·15	1·13	8·00	1·32	0·60	n.s.	0·53	n.s.
AV	8·28	1·72	7·33	1·44	6·80	1·55	1·73	n.s.	1·51	n.s.
VV	6·61	1·79	5·13	2·02	3·45	2·32	2·39	< 0·02	3·36	< 0·01
VA	5·79	2·11	5·53	1·27	3·83	1·19	0·93	n.s.	6·07	< 0·01

The results of this analysis thus indicate that the only change in the pattern of response between conditions in this test, comparing young (A) and middle-aged (B) groups, occurs in the condition where the subject has to note a number which is presented twice visually (*VV*). After the age of 60 years, however, the deterioration in retention of visual information is a great deal more pronounced. This occurs in all cases where the number is originally presented visually, regardless of the modality in which it is repeated (*VV* or *VA*). These results might be summarized as follows. The performance of groups of normal adults in this test confirms Buschke's finding that there is poorer retention of visual as opposed to auditory information. With increasing age this discrepancy between retention of information in the two modalities increases and the deterioration of the retention of visual information is particularly marked after the age of 60 years.

DISCUSSION

Before discussing the significance of these findings, we might first consider the possibility that the decline in the retention of visual information with age is due to a progressive falling off in visual acuity. As already stated, we took the precaution of ensuring that the visual material presented was of such a size and intensity as to be seen clearly by all subjects.

It is of course already a well-established finding that short-term memory tends to decline with age (Clay, 1954; Jerome, 1962; Welford, 1958, 1962). In his analysis of changes in human performance with age, Welford (1958) noted that the decline in short-term retention was one of the most important changes occurring in old age, and he and others have pointed out the profound and multiple effects which this may have on the mental output of older people. Other workers have noted that the decrement in short-term memory with age is particularly severe when subsequent activity intervenes during the period of retention. Cameron (1943) compared senile patients and young adults for their ability to retain digit sequences when the intervening period between retention and recall was filled by a spelling task. While both

groups coped adequately with the task when recall was preceded by a blank interval, the senile group's performance declined sharply when the intervening period was occupied by the spelling task. Kay (1953) and Kirchner (1958), using a visual retention test where the subject had to respond to sequences of flashing lights, showed that older subjects had great difficulty when they were required to respond to lights flashed one or more places earlier in the sequence. It seemed, that for older subjects, the very act of responding (in these experiments this involved pressing a key) interfered with short-term retention. Griew (1958) obtained similar results with age in a visual tracking task involving short-term retention of relevant signals. Speakman (1958), and more recently Broadbent and Heron (1962), have demonstrated that, when a task involving short-term retention is combined with a secondary distracting task, older subjects tend to perform badly in at least one of the tasks, whereas younger subjects are able to cope with both tasks at a reasonable level. Inglis (1960) presented sets of spoken digits dichotically to younger and older subjects, asking them to report first the digits received at one ear, before reporting those received at the other ear. Although there was little difference between the groups in recalling the set of digits which they were instructed to report first, the older subjects were much less successful in recalling the digits to be reported second. Fraser (1958) has also demonstrated that older subjects perform badly on a digit span test when the rate of presentation is slow, a condition which allows more opportunity for the effect of distraction on short-term memory.

Our main finding is the suggestion that the decline in immediate memory with age is much more severe for visual than for aural information. This is evident in the greater difficulty experienced by the older subjects when they were required to retain a visual signal, regardless of the modality of the later signal to which it had to be linked. Samples from each age-group were asked at the end of their test to state their own opinion as to whether they had been aware of depending on visual images in retaining the visual material, or whether they had been conscious of recoding into auditory signals before storing the information. Responses of the subjects indicated that around 40% felt that they were depending mainly on visual images, a further 40% were aware of translating these into auditory images, and the remaining 20% were doubtful about the over-all strategy which they had used. Reports by such workers as Sperling (1960) and Conrad (1962), however, suggest that the subjective comments of subjects on the strategies used in tasks of this nature are not very reliable. Errors made in the recall of visually presented letter sequences have indicated confusions of an auditory kind (e.g. *E* for *C*); confusions of a visual kind (e.g. *E* for *F*) have been relatively infrequent. The tendency to transfer visual into auditory information before storage would presumably exert a greater load on short-term retention, and this may partially account for the more rapid decline of visual retention in older subjects.

An alternative interpretation which might be suggested by these findings is that auditory and visual information are stored separately in short-term memory. Most investigations of short-term retention, particularly those which have examined changes with age, have concentrated on the retention of aurally presented information. It is of some interest that those investigations which have shown the most marked variation in performance with age have utilized visual displays (Kay, 1953;

Kirchner, 1958; Griew, 1958; Broadbent & Heron, 1962). There have been, however, some attempts to examine the properties of the visual after-image and its relation to short-term retention. The work of Averbach and Corriell (1961) and Mackworth (1963) indicates that visual images have a brief duration and tend to be erased by the arrival of visually new information. Crossman (1964) has produced evidence to suggest that auditory images last considerably longer and that they are relatively resistant to auditory erasure. It is obviously more important for the individual to have an efficient and less vulnerable auditory storage system as auditory information is always transient in its presentation, while visual information may usually be scanned for some time.

It would thus appear that there are at least two factors involved in the rapid decline of short-term visual retention with age. The unstable nature of the visual image may normally be partially compensated by recoding visual information into aural form, thus allowing it to be stored in the more stable and resistant auditory storage system. The apparent decline in the capacity of short-term storage with age suggests, however, that with older people, this recoding procedure leads to an overloading of an already relatively limited storage capacity. Welford (1958) has already dealt in some detail with the practical consequences of the diminution of short-term storage capacity and its vulnerability to interference in old age, and it seems clear that this decline in the efficiency of the brain helps to explain many of the difficulties shown by elderly people in situations demanding the rapid processing of incoming information. Our results may allow us to add the suggestion that the older person may be at a further disadvantage if the data to be handled by the short-term memory storage system are mainly visual, rather than aural. These effects may partially account for the slowing down in mental performance so characteristic of elderly people. They might also help to explain the marked variations in performance of older subjects between different types of psychological tests. It is now well established that the deterioration in intellectual efficiency in old age is not a uniform process. Some intellectual tests (e.g. Progressive Matrices) show a fairly rapid decline in performance in old age, as compared with some other tests which show comparatively little change with age. The work of Welford and others suggests that intellectual deterioration is an amalgam made up of many specific changes in mental performance which are not easily made evident by such crude measuring instruments as tests of general intelligence. Welford (1958) has already suggested that the difficulty experienced by the older person on such a test as the Progressive Matrices is one of effecting a spatial transposition of data not actually present in the test situation. Most of the standard intellectual tests are essentially visual and require short-term retention and manipulation of visual information for their execution (e.g. Progressive Matrices, Sorting Tests, Design Reproduction, Maze Tracking, Block Design). Rather than interpret the results of such tests as implying a general deterioration in reasoning capacity, it might be argued that they were at least partially due to a decline in short-term retention which directly affects the efficient processing of information. Finally, on the assumption that pathological deterioration may be an exaggeration of the changes occurring in normal ageing, it would seem that a test which differentiates between the visual and auditory components in immediate memory might offer a useful tool in investigating the effects of brain damage.

The work of two of the authors, J. C. and J. S. L., is supported by grants from the Medical Research Council.

References

Averbach, E. & Corriell, A. S. (1961). Short-term memory in vision. *Bell Syst. Tech. J.* **40**, 309–28.

Broadbent, D. E. & Heron, A. (1962). Effects of a subsidiary task on performance involving immediate memory by younger and older subjects. *Brit. J. Psychol.* **53**, 189–98.

Buschke, H. (1962). Auditory and visual interaction in immediate memory. *J. Psychiat. Res.* **1**, 229–39.

Cameron, D. E. (1943). Impairment of the retention phase of remembering. *Psychiat. Quart.* **17**, 395–404.

Chapman, J. A. & McGhie, A. (1962). A comparative study of disordered attention in schizophrenia. *J. ment. Sci.* **108**, 455, 487–500.

Clay, H. M. (1954). Changes of performance with age on tasks of varying complexity. *Brit. J. Psychol.* **14**, 7–13.

Conrad, R. (1962). An association between memory errors and errors due to acoustic masking of speech. *Nature, Lond.* **193**, 1314.

Crossman, E. R. F. W. (1964). Information processes in human skill. *Brit. med. Bull.* **20**, no. 1, 32–7.

Fraser, D. C. (1958). Delay of immediate memory with age. *Nature, Lond.*, **179**, 4564.

Griew, S. (1958). Age changes and information loss in performance of a pursuit tracking task involving interrupted preview. *J. exp. Psychol.* **55**, 486–91.

Inglis, J. (1960). Dichotic stimulation and memory disorder. *Nature, Lond.*, **186**, 181–2.

Jerome, E. A. (1962). In *Social and Psychological Aspects of Ageing*, ed. Tibbits, C. and Donahue, W. *Proc. Fifth Cong. Int. Ass. Gerontology*, vol. I. New York: Columbia University Press.

Kay, H. (1953). Experimental studies of adult learning. Ph.D. Thesis, University of Cambridge.

Kirchner, W. K. (1958). Age difference in short-term retention of rapidly changing information. *J. exp. Psychol.* **55**, 352.

Lawson, J. S., McGhie, A. & Chapman, J. (1964). Perception of speech in schizophrenia. *Brit. J. Psychiat.* **110**, 465.

McGhie, A. & Chapman, J. A. (1961). Disorders of attention and perception in early schizophrenia. *Brit. J. Med. Psychol.* **34**, 103–16.

Mackworth, J. F. (1963). The duration of the visual image. *Canad. J. Psychol.* **17**, 62–81.

Speakman, D. (1958). Cited by Welford, A. T. *Ageing and Human Skill*, p. 251. London: Oxford University Press.

Sperling, D. (1960). The information available in brief visual presentations. *Psychol. Monogr.* **74**, no. 11.

Welford, A. T. (1958). *Ageing and Human Skill*. London: Oxford University Press.

Welford, A. T. (1962). On changes of performance with age. *Lancet*, 17 February, 335–9.

(*Manuscript received* 14 *February* 1964)

Brit. J. Psychol. (1965), 56, 1, pp. 77–80
Printed in Great Britain

SOME CONFIRMATORY RESULTS ON AGE DIFFERENCES IN MEMORY FOR SIMULTANEOUS STIMULATION

By D. E. BROADBENT and MARGARET GREGORY
Medical Research Council, Applied Psychology Research Unit, Cambridge

Two experiments were carried out to investigate further the relationship between age and immediate memory for two streams of material applied simultaneously to two sensory channels. The material was presented over ordinary television transmitters during programmes, viewers being requested to send in their responses. In both experiments the viewers saw three items and heard another three. In Expt. I all items were digits; in Expt. II both the visual and the spoken items consisted of letters and digits, the viewers being requested to order their responses either channel by channel or class by class. Both experiments showed a deterioration in performance with age which started earlier when the task required rapid alternation between the senses. Expt. II also allowed comparison between the two modes of recall: the eye–ear mode resulted in much superior performance.

Inglis & Caird (1963) have reported relations between age and deterioration on a test of immediate memory for two streams of material applied simultaneously to the two ears. The present note reports some experiments designed to confirm the effect of age on memory for material presented to two senses. The technique employed was to deliver the stimuli over the ordinary television transmitters during programmes, and to appeal to viewers to send in their responses by post. This method gives weak control of the stimulus, and possibly some danger of sampling biases in the respondents. The population is, however, large and the findings may fairly be taken as complementary to small-scale laboratory studies: as will be seen, some internal controls are included.

A second purpose of one experiment was as follows. When six digits are rapidly delivered, three to one ear and three to the other, it is hard to respond to them in an order which alternates between ears. When, however, the six items contain three digits and three other words, with each ear receiving two items of one class and one of the other, it is as easy to reproduce the material class by class as it is ear by ear (Gray & Wedderburn, 1960; Yntema & Trask, 1963; Broadbent & Gregory, 1964). In view of other differences between performance using the two ears as separate sensory channels, and performance using the eye and ear (Broadbent & Gregory, 1961), it was desirable to repeat this experiment also using the eye and ear.

Method

Experiment I

This experiment was conducted in the middle of the evening over transmitters serving the East Anglia, Glasgow and Southampton regions. Viewers were presented with three tasks.

Task A. A series of six visual digits, each remaining on the screen for $\frac{1}{3}$ sec., and followed by an interval of clear screen for $\frac{1}{3}$ sec. before the next digit.

Task B. Three visual digits appearing in succession each for $\frac{1}{3}$ sec. and with a blank interval of $\frac{2}{3}$ sec. between each digit and the next. During each of the blank intervals, a spoken digit was transmitted over the sound channel.

Task C. A repetition of *B*, using different digits.

Between the three presentations there were pauses, in which viewers were asked to write down as many of the digits as they could remember. A form was provided in the weekly television magazine; and this also contained spaces for the age and sex of the viewer, as well as the ages of others in the same room. No practice was given; during the programme instructions were given to recall the visual material before the auditory, but this instruction was not repeated on the form and was often misunderstood. The analysis of results is therefore simply in terms of correct items regardless of order.

Experiment II

This experiment was conducted at 7.45 a.m. over the national network. Viewers were again presented with three tasks.

Task A. Six successive visual digits as in the previous experiment.

Task B. Three visual items consisting of a digit, a letter of the alphabet, and another digit, each presented for $\frac{1}{3}$ sec. with $\frac{1}{3}$ sec. interval. Simultaneous with each item there was an auditory item of the other class.

Task C. A repetition of *B* with different items.

As before viewers were asked to write their responses on a form which was provided. For task *B*, the instructions were to reproduce the visual items before the auditory; and for task *C* to reproduce first the letters and then the digits. These instructions were given on the form as well as in the programme. Viewers had previously received a demonstration similar to Expt. I, which could be regarded as practice.

RESULTS

The number of respondents who performed each task correctly and incorrectly is shown in Table 1. Correct responses as a percentage of total responses are shown in Figs. 1 and 2.

Table 1. *Number of subjects of each sex in each age-group making correct and incorrect responses*

Conditions		Under 16 years		16–25		26–35		36–45		Over 45 years	
		Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Simple spans Expt. I	Right	103	201	226	379	219	314	164	315	193	343
	Wrong	9	10	9	31	7	11	10	19	5	25
Expt. II	Right	81	29	224	114	146	95	160	135	152	153
	Wrong	6	3	12	12	11	7	12	13	11	28
Bisensory spans Expt. I	Right*	90	177	217	367	197	287	154	284	163	303
	Wrong	22	36	19	44	29	40	20	50	35	67
Expt. II Eye–ear	Right	34	12	123	48	73	39	87	58	53	48
	Wrong	53	20	113	79	84	66	85	90	110	133
Letters–digits	Right	25	14	77	43	31	19	35	17	18	11
	Wrong	62	18	159	84	126	86	137	131	145	170

* Tasks *B* and *C* both correct.

In Expt. I a correct response had to have the correct items regardless of order; likewise in Expt. II, except that the two halves of the list had to be kept separate. The numbers of subjects are not identical for simple spans and for bisensory spans, because some viewers completed only one or other of the tasks, adding a note to say that a broken pencil or a similar accident had held them up.

(*a*) *Little difference between ages in simple visual span.* In neither experiment was χ^2 significant overall between the five age-groups. Picking out the 16–25 and over-45 groups, neither sex alone showed a significant difference in Expt. I, and in Expt. II only the females did so ($P < 0{\cdot}05$). In both experiments the males showed some

signs of superiority over females of the same age: in Expt. I, at the over-45 level: and, in Expt. II, at the 16–25, 36–45 and over-45 levels there were differences signi-

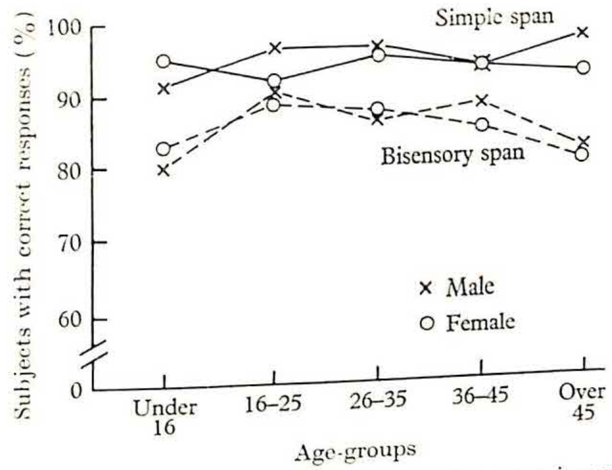


Fig. 1. Performance in Exp. I for simple spans involving six successive visual items, and for bisensory spans involving three visual and three auditory items presented in rapid alternation.

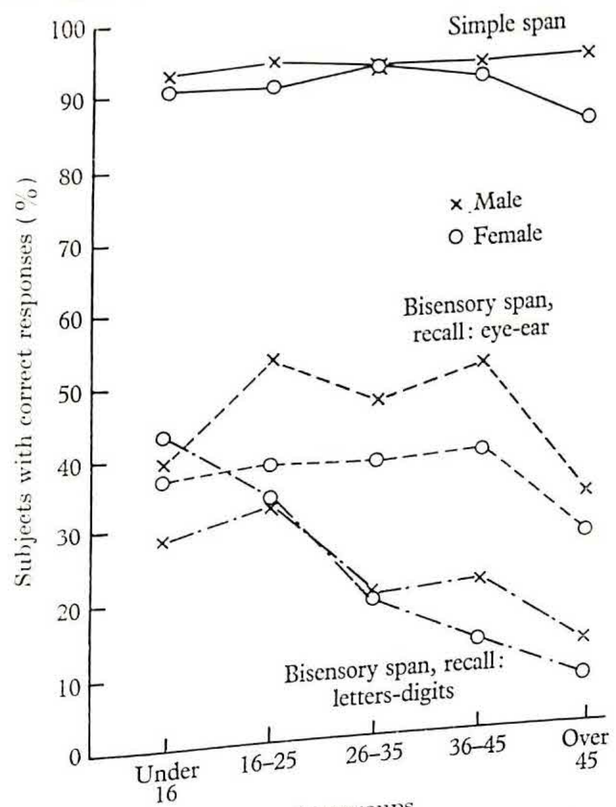


Fig. 2. Performance in Exp. II for simple spans involving six successive visual digits, and for bisensory spans involving three visual and three simultaneous auditory items. In the latter case the visual items were a digit, a letter, and a digit while the auditory items were a letter, a digit, and a letter. Recall was organised either by sensory channel or by class of material.

ficant at the 0·05 level or better. These latter differences are of course likely to be due to sampling factors: it is the absence of an effect of age which is important, because it suggests that the differences in the next section are not due to an older sample of inferior general ability.

(*b*) *Drop with age in all bisensory conditions.* The overall χ^2 between age-groups was significant for each experiment for the bisensory conditions. In Expt. I, for tasks *B* and *C* combined, the over-45 group was worse than the 16–25 both for males alone and for females alone ($P < 0{\cdot}01$). In Expt. II, task *B* (eye first then ear) gave a similar difference ($P < 0{\cdot}01$ for each sex). Task *C* (letters first then digits) not only gave the same result, but even gave significant differences between the 16–25 group and the 25–35 group for each sex ($P < 0{\cdot}05$). This task requires alternation between one sense and the other, and it will be recalled that in Expt. I some subjects misunderstood the instructions and responded in that way. The percentage of such correct responses in the alternating order decreased with age, the difference between ages 16–25 and over-45 being significant both for males and for females ($P < 0{\cdot}05$). These results applied to viewers who were alone as well as to those in company.

(*c*) *Difference between the eye–ear and digit–letter modes of recall.* In Expt. II tasks *B* and *C* give widely different performances, task *C* being harder. The difference was found for all ages and both sexes and is of course very highly significant indeed. Thus the equal status of the two methods of grouping, found by Gray and Wedderburn when the two sensory channels were the two ears, does not apply when they are the eye and ear.

(*d*) *Abnormality of performance below age* 16. It was unexpected that children under 16 performed as well as the 16–25 group at task *A*, the simple span, in both experiments. Presumably an unusually intelligent sample was selected by the nature of the programmes. It is notable even so that, in Expt. I, both boys and girls under 16 did worse at the bisensory task than did the 16–25 group ($P < 0{\cdot}05$ in each case). In Expt. II the same is true for boys for task *B*, but not for girls; and in task *C* both boys and girls did unusually well compared with their elders. Tentatively one may suggest that pre-adults have difficulty in organizing a bisensory situation but perhaps alternate very rapidly between different channels. This point suggests further investigation of a laboratory type.

Thanks are due to many persons whose co-operation was vital to these experiments, but especially to Anglia Television, from whose studios the programmes were produced, and to George Noordhof, who produced them.

REFERENCES

BROADBENT, D. E. & GREGORY, M. (1961). On the recall of stimuli presented alternately to two sense-organs. *Quart. J. Exp. Psychol.* **13**, 103–9.

BROADBENT, D. E. & GREGORY, M. (1964). Stimulus set and response set: the alternation of attention. *Quart. J. Exp. Psychol.* **16**, 309–12.

GRAY, J. A. & WEDDERBURN, A. A. (1960). Grouping strategies with simultaneous stimuli. *Quart. J. Exp. Psychol.* **12**, 180–4.

INGLIS, J. & CAIRD, W. K. (1963). Age differences in successive responses to simultaneous stimulation. *Canad. J. Psychol.* **17**, 98–105.

YNTEMA, D. B. & TRASK, F. P. (1963). Recall as a search process. *J. Verb. Learn. Verb. Behav.* **2**, 67–74.

(*Manuscript received* 24 *June* 1964)

Brit. J. Psychol. (1965), 56, 1, pp. 81–86
Printed in Great Britain

THE TRANSFER OF POSITIVE AND NEGATIVE LEARNING BY NORMAL AND SEVERELY SUBNORMAL CHILDREN

By P. E. BRYANT

Medical Research Council Social Psychiatry Research Unit, Institute of Psychiatry, Maudsley Hospital, London, S.E. 5

An experiment was designed in which all the subjects learned the same initial task; half then learned a transfer task to which positive learning could be transferred, while the other half learned one to which negative learning could be transferred. It was found that severely subnormal subjects transferred negative learning more easily than positive learning, while normals showed the opposite pattern. It is suggested that this was due to the fact that the positive learning was specific and the negative learning relatively general.

The aim of this experiment was to investigate the possibility that normal and severely subnormal subjects depend on different learning processes in transfer situations. This possibility was suggested by the results of a previous study (Bryant, 1964), which dealt with the effects of verbal instruction on the transfer of a conceptual strategy. Severely subnormal and normal children learned two consecutive tasks in both of which they had to respond on the basis of the dimension of number and to ignore the dimension of size. During the initial task half of the subjects were instructed verbally while the other half were not. None of the subjects was instructed in the transfer task. Instruction improved the performance of both the normal and the severely subnormal subjects at the time it was given in the initial task. This instruction was also followed by improved transfer task performance in the normal group. However, in the severely subnormal group those subjects who had previously been instructed learned the transfer task more slowly than the subjects who had received no instruction.

One possible explanation for this difference between the two groups is based on a distinction between positive and negative learning. It was suggested that both tasks involved learning not only of the correct dimension (positive learning), but also of the incorrect dimension (negative learning). The verbal instruction given in the initial task concerned only the correct dimension, and it was therefore suggested that this instruction strengthened positive learning only. Since this instruction was followed by good transfer with normals and poor transfer with subnormal subjects, it was possible that the subnormal group depended less on previous positive learning and more on previous negative learning in transfer situations.

It seemed that this hypothesis could easily be tested by designing an experiment in which there were two transfer tasks, one to which positive learning could be transferred and one to which negative learning could be transferred. It was predicted that severely subnormal subjects would transfer to the negative transfer task more readily than normal subjects, and that the normal subjects would transfer to the positive transfer task more readily than the severely subnormal subjects.

Accordingly, an experiment was carried out in which severely subnormal and normal children all learned an initial task in which the correct dimension was colour

and the incorrect dimension size. Thereafter half the subjects in each group learned the positive transfer task, in which the correct dimension was again colour, and the other half learned the negative transfer task, in which the incorrect dimension was again size.

Method

Subjects

There were eighty subjects. Forty were severely subnormal children who were patients in an institution for the subnormal, and the other forty were normal children attending a local junior school. The subjects were divided into four groups of twenty subjects: a subnormal experimental group, a normal experimental group, a subnormal control group and a normal control group. The groups were matched in terms of the average M.A. of the subnormal groups and the average C.A. of the normal groups. The mental and chronological ages are given in Table 1.

Table 1. *Mental and chronological ages*

Subnormal groups	Av. M.A.	Range
Experimental	5 y. 8 m.	5 y–6 y. 8 m.
Control	5 y. 6 m.	4 y. 7 m.–6 y. 10 m.
Normal groups	Av. C.A.	
Experimental	5 y. 10 m.	5 y. 6 m.–6 y. 2 m.
Control	5 y. 11 m.	5 y. 7 m.–6 y. 2 m.

Procedure

There were three tasks: the initial task, the positive transfer task and the negative transfer task. These tasks were designed in such a way that only positive learning could be transferred from the initial task to the positivé transfer task and only negative learning could be transferred from the initial task to the negative transfer task. For each task cards were used which differed in terms of two dimensions. The subject had to learn to respond on the basis of one dimension, the correct dimension (positive learning), and to avoid responding on the basis of the other dimension, the incorrect dimension (negative learning). Positive learning could be transferred to the positive transfer task because it had the same correct dimension as the initial task. Negative learning could be transferred to the negative transfer task because it had the same incorrect dimension as the initial task. Different sets of sixteen cards were used for each task. Each card consisted of a capital letter on a square white background. The letters varied in size and colour. Cards were presented four at a time and the subject had to learn to select the correct card on each trial. Each card was placed over a shallow saucer, and on every trial a small sweet was concealed in the saucer which was covered by the correct card. In this way the subjects were rewarded whenever they made a correct choice.

In the initial task all sixteen cards had the same letter, H, on them, but varied in terms of the colour and size of this letter. There were four colours and four sizes. Cards were presented four at a time in such a way that all the four colours and all the four sizes were represented in each trial; the combinations of colours and sizes were varied systematically from trial to trial (Table 2*a*). The subject had to learn to respond to a particular colour and was rewarded each time it was chosen; e.g. he might have to learn to respond to the colour green at each trial, irrespective of size. The correct dimension was therefore colour and the incorrect dimension was size.

In the positive transfer task the correct dimension was again colour, but the colours were different from those used in the initial task. The incorrect dimension was letter. The sixteen cards were all 5 in. square; four different letters of similar size and four different colours were used. Table 2*b* gives the sixteen cards as they were presented in the first four trials. As in the initial task, the subject again had to learn to respond to a particular colour which was rewarded each time it was chosen; e.g. he might have to respond to the colour orange at each trial. However, in this task he had to learn that the particular letter on each card was irrelevant.

In the negative transfer task the incorrect dimension was again size, although the sizes were different from those used in the initial task. The correct dimension was letter. The sixteen cards were formed from the combinations of four different letters and four different sizes. All the letters were black. Table 2*c* shows how the cards were presented in the first four trials. The

Table 2

(a) First four trials in the initial task (letter—H)

Positions...	Extreme left (in.)	Centre left (in.)	Centre right (in.)	Extreme right (in.)
Trial 1	9 × 9, green	7 × 7, yellow	5 × 5, red	3 × 3, grey
Trial 2	5 × 5, yellow	3 × 3, green	9 × 9, grey	7 × 7, red
Trial 3	3 × 3, red	5 × 5, grey	7 × 7, green	9 × 9, yellow
Trial 4	7 × 7, grey	9 × 9, red	3 × 3, yellow	5 × 5, green

(b) First four trials in the positive transfer task (size—5 in. square)

Positions...	Extreme left	Centre left	Centre right	Extreme right
Trial 1	Blue L	Brown X	Black V	Orange T
Trial 2	Black X	Orange L	Blue T	Brown V
Trial 3	Orange V	Black T	Brown L	Blue X
Trial 4	Brown T	Blue V	Orange X	Black L

(c) First four trials in the negative transfer task (colour—black)

Positions...	Extreme left (in.)	Centre left (in.)	Centre right (in.)	Extreme right (in.)
Trial 1	2 × 2 L	4 × 4 X	6 × 6 V	8 × 8 T
Trial 2	6 × 6 X	8 × 8 L	2 × 2 T	4 × 4 V
Trial 3	8 × 8 V	6 × 6 T	4 × 4 L	2 × 2 X
Trial 4	4 × 4 T	2 × 2 V	8 × 8 X	6 × 6 L

subject had to learn to respond to a particular letter which was rewarded each time it was chosen; e.g. he might have to respond to the letter V at each trial. As in the initial task, he had to learn not to respond on the basis of size.

The dimensions which were correct and incorrect in the three tasks are summarized in Table 3. For each task learning was to a criterion of eight successive correct trials. In the control groups, half the subjects learned only the positive transfer task, while the other half learned only the negative transfer task. These groups were introduced to check that the two transfer tasks were matched for level of difficulty. In the experimental groups, immediately after the subjects had learned the initial task, half of them went on to learn the positive transfer task and the other half the negative transfer task. Hence, a relatively good performance in the positive transfer task and a poor performance in the negative transfer task by an experimental group would indicate that this group transferred on the basis of previous learning of the correct dimension or, in other words, on the basis of previous positive learning. Similarly, the reverse pattern would indicate that this group depended more on previous negative learning in this kind of transfer situation.

Table 3. *Correct and incorrect dimensions in the three tasks*

	Correct dimension	Incorrect dimension
Initial task	Colour	Size
Positive transfer task	Colour	Letter
Negative transfer task	Letter	Size

Results

Each subject was scored in terms of the number of trials between the first correct choice and the beginning of the eight criterion trials.

Control groups

In the severely subnormal control group the average number of trials to reach criterion was 14·8 for the subjects who learned the positive transfer task and 15·3

for those who learned the negative transfer task ($t = 0.75$, D.F. $= 18$). The corresponding averages in the normal control group were 11·4 and 12·1 ($t = 0.87$, D.F. $= 18$). Hence, these results indicated that there was no difference between the two tasks in terms of level of difficulty, either for the severely subnormal or for the normal subjects.

Experimental groups

The initial task and the transfer task scores were analysed separately in two analyses of variance, in which *Groups* (severely subnormal and normal subjects) and *Conditions* (subjects who learned positive and negative transfer tasks) were compared.

The normal group learned the initial task more rapidly than the severely subnormal group (Table 4). The average number of trials to reach criterion was 17·9 for the severely subnormal group and 12·8 for the normal group ($P < 0.025$). The lack of a significant difference between Conditions or Groups × Conditions interaction showed that there was no difference in the learning of this task between the subjects who later learned the positive transfer task and those who later learned the negative transfer task.

The transfer task scores and analysis are given in Table 5, which shows a signi-

Table 4. *Analysis of initial task scores*

Scores

	n	Av. trials to criterion	S.D.
Experimental subnormal group	20	17·9	5·12
Experimental normal group	20	12·8	5·8

Analysis of variance

Source	Sums of squares	D.F.	F	Level of significance
Groups	260·1	1	6·041	< 0·025
Conditions	0·9	1	0·021	N.S.
Groups × conditions	8·1	1	0·188	N.S.

Table 5. *Analysis of transfer task scores*

Scores

	Positive transfer task			Negative transfer task		
	n	Av. trials to criterion	S.D.	n	Av. trials to criterion	S.D.
Experimental subnormal group	10	9·2	4·85	10	5·3	3·55
Experimental normal group	10	5	3·56	10	7·8	4·52

Analysis of variance

Source	Sums of squares	D.F.	F	P
Groups	7·225	1	—	—
Conditions	3·025	1	—	—
Groups × conditions	112·225	1	4·32	< 0·05

ficant Groups × Conditions interaction ($P < 0{\cdot}05$). The severely subnormal subjects learned the negative transfer task more rapidly than the positive transfer task. For this group the average number of trials was 5·3 in the negative transfer task and 9·2 in the positive transfer task. On the other hand, the normal group learned the positive transfer task more rapidly than the negative transfer task, the average number of trials to criterion being 5·0 in the positive transfer task and 7·8 in the negative transfer task.

The scores of the two control groups had already shown that there was no difference between the two tasks in terms of level of difficulty for either severely subnormal or normal subjects. It can therefore be concluded that the significant Groups × Conditions interaction shows that previous experience of the initial task had a different effect on the two groups in the transfer situation. Severely subnormal subjects transferred more effectively to the transfer task which had the same kind of negative learning as the initial task. Normal subjects transferred more effectively to the transfer task which had the same kind of positive learning as the initial task. This difference between the two groups had been predicted.

Discussion

In this experiment severely subnormal subjects transferred previous negative learning more effectively than previous positive learning. Normal subjects showed the opposite pattern. These results imply that there is a qualitative difference between the two groups in the learning processes on which they depend in this kind of transfer situation. Since the ability to transfer is basic to behaviour which is on a conceptual level, such a qualitative difference is obviously relevant to the study of cognition in mental defect.

There seem to be two possible explanations for this difference between the two groups. The first is that severely subnormal subjects learn more effectively from non-reinforced trials than from reinforced trials, and for this reason transfer learning of what to ignore more readily than learning of what to approach. However, experiments by House & Zeaman (1958) and House, Orlando & Zeaman (1957) imply that this is not so, at least in a simple object discrimination situation. The second possibility is that in all the tasks the learning of the correct dimension was confined to only one aspect of that dimension. In the initial task, for example, the subjects had to learn to respond to one colour only, but they also had to learn to ignore all the sizes, since the correct colour was represented in all the four sizes. Thus the learning of the correct dimension, the positive learning, could be said to be relatively specific since it concerned only one aspect of that dimension. Similarly, the learning of the incorrect dimension, the negative learning, could be said to be relatively general since it concerned all four aspects of that dimension.

It is therefore possible that what has been found is that normal subjects are able to transfer on the basis of learning which is specific, while severely subnormal subjects are less able to do this and, in order to generalize learning to a new situation, have to fall back on learning which is already relatively general. This is a tentative hypothesis, but if it were supported by further experimental evidence it could indicate a major difference in cognitive processes between severely subnormal and normal children.

This experiment formed part of a thesis submitted to the University of London in partial fulfilment of the requirements for the Ph.D. degree. The author is indebted to Dr B. M. Hermelin and Dr N. O'Connor, his supervisors, and to the staff and children of St Lawrence's and Botleys Park Hospitals, where this research was carried out.

REFERENCES

BRYANT, P. E. (1964). The effect of a verbal instruction on transfer in normal and severely subnormal children. *J. Ment. Def. Res.* **8**, 35–43.

HOUSE, B., ORLANDO, R. & ZEAMAN, D. (1957). The role of positive and negative cues in the discrimination learning of mental defectives. *Percept. Motor Skills*, **7**, 73–9.

HOUSE, B. & ZEAMAN, D. (1958). Reward and non-reward in the discrimination learning of imbeciles. *J. Comp. Physiol. Psychol.* **56**, 614–18.

(*Manuscript received* 12 *March* 1964)

Brit. J. Psychol. (1965), 56, 1, *pp.* 87–93
Printed in Great Britain

THE EFFECT OF ATTENTION ON THE SLOPE OF GENERALIZATION GRADIENTS

By N. J. MACKINTOSH

Institute of Experimental Psychology, University of Oxford

Forty-eight rats were given 100 rewarded trials with a single rectangle, and were then tested for generalization between this first rectangle and another differing from it either in brightness or in orientation. Thirty-two of the rats had received prior training either on a successive brightness or on orientation discrimination. The sixteen subjects who had received no prior training showed intermediate gradients of generalization. Of the pretrained subjects, those who were tested for generalization along the same dimension as their pretraining discrimination showed significantly steeper gradients, while those who were tested for generalization along the opposite dimension showed significantly less steep gradients.

On one interpretation, the controversy between Lashley & Wade (1946) and Hull (1947) concerning stimulus generalization centres round the role of attention in determining gradients of generalization. Lashley & Wade claimed that a gradient of generalization is only formed after experience with a single stimulus, if during this experience the subject attended to the dimension along which generalization is subsequently measured, i.e. classified the stimulus in terms of that dimension. Hull, on the other hand, claimed that gradients of generalization are produced automatically. This point has been obscured by Lashley & Wade's further claim that animals will only attend to a particular dimension if they are given discrimination training between two points on the dimension. This strong claim, although apparently substantiated by results briefly reported by Lashley & Wade, has since then been seriously discredited: experiments by Spence (quoted by Hull, 1947) and Grice (1948, 1951) with rats, and Warren & Hall (1956) and Warren & Brookshire (1959) with monkeys, have all failed to confirm Lashley & Wade's results.

These later results have been thought to refute Lashley & Wade's position completely, but if the above analysis is correct this is clearly untrue: all that they show is that experience with two points along a dimension is not a necessary condition for attending to that dimension, and that during experience with a single stimulus animals will classify the stimulus as occupying a certain region along at least one dimension (if not more, see Butter, 1963). The weaker claim, that such a classification is affected by attention, remains untouched. Furthermore, such evidence as there is tends to support it: McCaslin, Wodinsky & Bitterman (1952) with rats, and Gentry, Overall & Brown (1958) with monkeys have shown that, under certain (rather ill-defined) circumstances, subjects who have been pretrained to select S' presented alone will not necessarily prefer S' to S'' in a subsequent test.

Previous experiments, however, have not attempted systematically to control the stimulus dimension to which an animal will attend, and to measure the effect of such a procedure on subsequent generalization testing. In the present experiment, three major groups of rats were trained to approach a singly presented stimulus—for example a White Horizontal (WH) rectangle. Generalization was measured by pairing this stimulus with a new one, and training animals either to continue to approach the

WH rectangle (Non-reversal), or to approach the new stimulus (Reversal): the greater the difference between rates of learning for the reversal and non-reversal discriminations, the steeper the gradient of generalization. In this way, generalization was measured along either the dimension of brightness—the WH rectangle was paired with a Black Horizontal (BH) rectangle, or along the dimension of orientation—the WH rectangle was paired with a White Vertical (WV) rectangle.

A control group was simply trained in the manner outlined above: they were not given any training designed to ensure that they would attend to any particular feature of the single WH rectangle. Two further groups (Relevant and Irrelevant) were pretrained on either a successive brightness or a successive orientation discrimination. This was designed to ensure that they would classify the single WH rectangle either in terms of its brightness or in terms of its orientation. The hypothesis to be tested was that if the WH rectangle was paired in the generalization test with a stimulus differing along the *same* dimension as had been used in successive pretraining (Relevant group) then a steeper gradient of generalization would result; whereas if the WH rectangle was paired with a stimulus differing along the dimension which had *not* been used in successive pretraining (Irrelevant group) then a shallower gradient would result.

METHOD

Subjects

The subjects were forty-eight experimentally naïve male hooded rats from the colony maintained at the Institute of Experimental Psychology, Oxford. They were 3–4 months old at the beginning of the experiment.

Apparatus

The apparatus was a modified Lashley jumping stand, with landing platforms below each window, and has been described in more detail elsewhere (Mackintosh, 1963). The stimuli were cut from $\frac{1}{4}$ in. black, white or grey Perspex, and mounted on 15×15 cm. brown hardboard doors. During single stimulus training, a grey wooden partition, 12 in. high, was clamped on to the jumping platform. It extended forwards as far as the partition dividing the two stimulus windows, and thus divided the apparatus into two single-window stands. A rat placed on one side of the jumping platform had no option but to jump to the single window in front of it.

Procedure

One week before the start of the experiment, the subjects were put on a feeding schedule of $1\frac{1}{2}$ hr. food per day. Water was always available in the home cages. Subjects were pretrained in the apparatus, to jump first to open and then to closed blank doors. During training, ten non-correction trials were given daily, with a 6–8 min. inter-trial interval. Correct responses were rewarded with 20 sec. access to food on the feeding platform behind the stimulus windows; no punishment was given for incorrect responses. Subjects were detained for 10 sec. on the landing platform in front of the locked negative door. Further details are given in Mackintosh (1963).

Experimental design

Stage 1: successive discrimination training. Thirty-two subjects (the two experimental groups) were given 200 trials of successive discrimination training. Half were trained to jump right to two black squares and left to two white squares; half were trained to jump right to two horizontal grey rectangles and left to two vertical grey rectangles. The size of the squares was $6{\cdot}5 \times 6{\cdot}5$ cm., and of the rectangles $10{\cdot}5 \times 4$ cm. The control group of sixteen subjects received no training at this stage.

Stage 2: single stimulus training. All subjects were given 100 rewarded jumps to a single stimulus. The side of the jumping platform on which they were placed (and hence the window to which they jumped) was varied from trial to trial in a random order. The sixteen subjects in each group (two experimental and one control) were divided into four equal subgroups; each subgroup of four rats was trained to jump to one of the following four stimuli: WH, WV, BH, BV. Each of these rectangles measured 10·5 × 4 cm.

Stage 3: non-reversal or reversal training. All subjects learned a simultaneous discrimination between the rectangle used in single stimulus training and another rectangle differing from it *either* in brightness *or* in orientation. The criterion of learning was 18 correct responses out of any 20 consecutive trials, with the last 10 all correct. For half the subjects in each group, their single stimulus was positive (non-reversal); for half, their single stimulus was negative (reversal).

In summary, the experimental design comprised three major groups (cf. Table 1).

Table 1. *Experimental design: progressive subdivision of each of the three main groups of rats over the three stages of the experiment*

	Group *C*. Control—no training in stage 1 (n = 16)	Group *R*. Stimulus dimension of stage 1 *relevant* in stage 3 (n = 16)	Group *I*. Stimulus dimension of stage 1 *irrelevant* in stage 3 (n = 16)
Stage 1 (200 trials) Preliminary training in successive discrimination Animals/subgroup	(None) (8) (8)	B/W 8; H/V 8	B/W 8; H/V 8
Stage 2 (100 trials) Single stimulus training Animals/subgroup	WH 2; WV 2; BH 2; BV 2		
Stage 3 (trials to criterion) Simultaneous discrimination training Animals/subgroup	Non-reversal 1; Reversal 1		

(i) *Group C*. The control group was given 100 single stimulus trials; eight then learned a non-reversal, and eight learned a reversal. Two of the subjects in each group of eight had been given one of the four possible single stimuli, and for one of these two the final discrimination was a brightness discrimination, for the other an orientation discrimination.

(ii) *Group R*. The relevant group was given 200 successive training trials: eight subjects on a brightness discrimination, eight on an orientation discrimination. In single stimulus training, two from each of these eight were trained with one of the four possible rectangles. Finally one member of each of these pairs learned a non-reversal, the other a reversal. For all subjects in this group the final discrimination involved the same stimulus dimension as their original successive training: if that had been a brightness discrimination, so was their final problem.

(iii) *Group I*. The irrelevant group was treated and subdivided exactly as group *R*, except that for this group the final discrimination always involved the opposite dimension from their successive discrimination: if that had been a brightness discrimination, their final problem was an orientation discrimination.

Results

Successive discrimination training

The brightness and orientation discriminations were not equally easy. Although all subjects reached a criterion of 80% correct responses over two successive days, the number of days to reach this criterion was 12·87 for subjects trained on the brightness discrimination, and 16·19 for those trained on the orientation discrimination

($t = 2{\cdot}70$: D.F. 30; $P < 0{\cdot}02$). Over the last twenty trials of training, they averaged 92·81% and 85·62% correct, respectively.

Non-reversal or reversal training

The results for this stage of the experiment are shown in Figs. 1 and 2. Fig. 1 gives the percentage of correct responses over trials 1–20 of training, and Fig. 2 gives the number of trials to criterion.

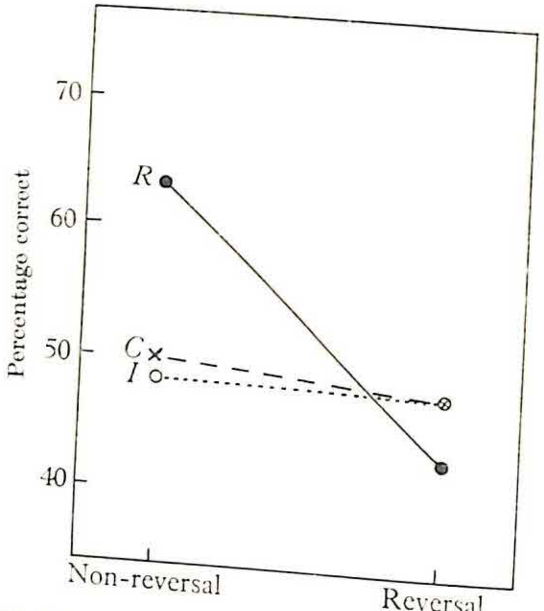


Fig. 1. Percentage correct responses over trials 1–20 of simultaneous discrimination training (stage 3) for non-reversal and reversal subgroups of groups C (control), R (relevant) and I (irrelevant); $n = 8$ per subgroup. The three main groups are distinguished in terms of the absence, relevance or irrelevance of preliminary training in successive discrimination at stage 1 (cf. Table 1).

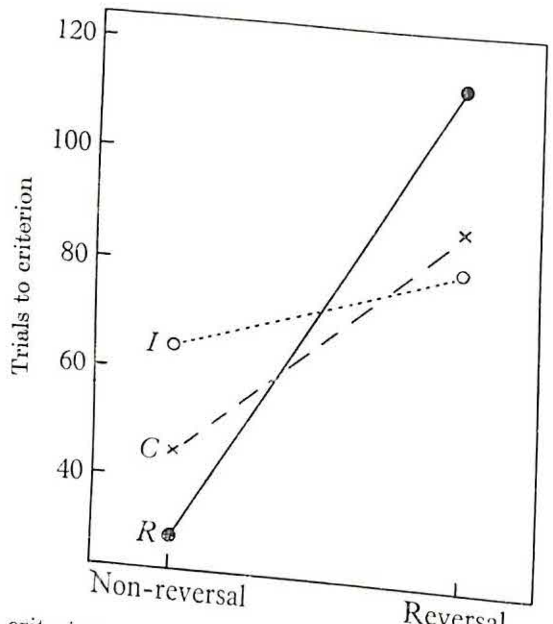


Fig. 2. Mean trials to criterion in simultaneous discrimination training (stage 3) for non-reversal and reversal subgroups of groups C, R and I; $n = 8$ per subgroup.

It is apparent from Fig. 1 that only group R showed a marked tendency during the first twenty trials of training to select the stimulus shown in single stimulus training. For both of the other groups the difference between non-reversal and reversal subjects was small. This impression is confirmed by the results of analyses of variance. In group R the difference between non-reversal and reversal subjects was highly significant ($F = 12{\cdot}19$; D.F. 1, 12; $P < 0{\cdot}01$); in group C, it fell short of significance ($F = 3{\cdot}86$; $0{\cdot}10 > P > 0{\cdot}05$); and in group I it did not even approach significance ($F < 1$).

Similar analyses were carried out on the trials to criterion scores (cf. Table 2), both for the main groups, and separately for the subgroups trained at this stage either on a brightness discrimination or on an orientation discrimination. In the case of group R, all differences were significant: for the BW subgroup, $F = 14{\cdot}06$, D.F. 1, 6, $P < 0{\cdot}01$; for the HV subgroup $F = 12{\cdot}76$; $P < 0{\cdot}01$; for the combined group, $F = 26{\cdot}76$; D.F. 1, 12; $P < 0{\cdot}001$. In group C, two out of the three differences were significant: for the BW subgroup, $F = 5{\cdot}99$, $P < 0{\cdot}05$; for the HV subgroup, $F = 3{\cdot}76$, $P > 0{\cdot}10$; for the combined group, $F = 9{\cdot}64$, $P < 0{\cdot}01$. In group I none of the differences approached significance (in all three cases, $F < 1$).

Table 2. *Trials to criterion for simultaneous discrimination training (stage 3) made by the animals in the three main groups in relation to the two types of discrimination, and both combined, with and without reversal; $n = 4$ for each of the subgroups shown*

Kind of simultaneous discrimination training in stage 3	Group C				Group R				Group I			
	Non-reversal		Reversal		Non-reversal		Reversal		Non-reversal		Reversal	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Brightness—BW	46·50	8·62	105·50	40·87	22·25	6·18	108·50	39·37	61·50	33·85	81·50	20·25
Orientation—HV	44·75	12·34	74·25	23·27	37·00	15·78	122·00	38·08	68·75	28·63	83·75	37·84
BW and HV combined	45·62	10·68	89·87	40·95	29·62	14·07	115·25	39·31	65·12	31·56	82·62	30·36

Thus far the results show that the difference in performance between non-reversal and reversal subjects was invariably highly significant in group R, usually significant in group C, and far short of significance in group I. This is precisely what would be expected on the hypothesis outlined above, but it is important to test a further point, namely whether this difference was significantly larger in group R than in group C, and larger in group C than in group I. As a more direct measure of this, the following procedure was adopted. Within each non-reversal and reversal subgroup, there were eight subjects, half trained on a brightness discrimination, half trained on an orientation discrimination, and each of these four subjects was trained with a different positive stimulus. The score of each subject in a non-reversal subgroup was subtracted from the score of the corresponding subject in a reversal subgroup (cf. last line of Table 1): these scores represent non-reversal–reversal differences, and two analyses of variance were performed on them. The first compared the magnitude of these differences in group R and group C: they were significantly larger in group R ($F = 23{\cdot}56$; D.F. 1, 9; $P < 0{\cdot}001$); the second compared group C with group I: the differences in group C were now significantly larger ($F = 8{\cdot}50$; $P < 0{\cdot}025$).

DISCUSSION

The results show unambiguously that the slope of a generalization gradient following single stimulus training is influenced by prior training. Naïve subjects, given experience of a single WH rectangle, show gradients of generalization along both brightness and (less significantly) orientation dimensions. In other words they tend to classify the WH rectangle as both white and (less markedly) horizontal. Thus Lashley and Wade's strong claim, that such classification only occurs if subjects have had experience with two points along a stimulus dimension, again appears to be refuted. However, their weaker claim that the occurrence of such classification depends upon attention receives convincing support: if an animal is given prior training on a successive brightness (or orientation) discrimination, this will increase its tendency to classify the WH rectangle as white (or horizontal), and correspondingly decrease its tendency to classify it as horizontal (or white).

In the case of group *C*, the differences between non-reversal and reversal subjects were relatively small: they were not significant over trials 1–20, and were only significant on trials to criterion for the BW subgroup. A partial reason for this may well be that the single stimulus training procedure did not demand any choice response from the animals, and that therefore they failed to look at the stimulus carefully (failed to establish adequate 'orienting responses'). This might be one reason why group *R* showed very much steeper generalization gradients; but it is important to note that it can certainly not be the only reason, since, if it were, group *I*, who also received training ensuring appropriate orientation towards the single stimulus, should also have shown steeper gradients.

The mechanism of attention implied by these results must in fact be an internal one, i.e. of the type suggested by Sutherland (1959). These results, therefore, provide additional support for such a theory of discrimination learning, and in particular for the implication that, the greater the tendency to attend to one stimulus dimension, the less will be the tendency to attend to another (Sutherland & Mackintosh, 1964).

This work forms part of a project on Stimulus Analysing Mechanisms, supported partly by the U.S. Office of Naval Research, Contract No. N 62558-3927 and partly by the Nuffield Foundation.

REFERENCES

BUTTER, C. M. (1963). Stimulus generalization along one and two dimensions in pigeons. *J. Exp. Psychol.* **65**, 339–46.

GENTRY, G. V., OVERALL, J. E. & BROWN, W. L. (1958). Discrimination after pretraining on component stimuli. *J. Comp. Physiol. Psychol.* **51**, 464–66.

GRICE, G. R. (1948). The acquisition of a visual discrimination habit following response to a single stimulus. *J. Exp. Psychol.* **38**, 633–42.

GRICE, G. R. (1951). The acquisition of a visual discrimination habit following extinction of response to one stimulus. *J. Comp. Physiol. Psychol.* **44**, 149–53.

HULL, C. L. (1947). The problem of primary stimulus generalization. *Psychol. Rev.* **54**, 120–34.

LASHLEY, K. S. & WADE, M. (1946). The Pavlovian theory of generalization. *Psychol. Rev.* **53**, 72–87.

MCCASLIN, E. F., WODINSKY, J. & BITTERMAN, M. E. (1952). Stimulus-generalization as a function of prior training. *Amer. J. Psychol.* **65**, 1–15.

Mackintosh, N. J. (1963). The effect of irrelevant cues on reversal learning in the rat. *Brit. J. Psychol.* **54**, 127–34.
Sutherland, N. S. (1959). Stimulus analysing mechanisms. In *Proceedings of a Symposium on the Mechanisation of Thought Processes*, vol. II, pp. 575–609. London: Her Majesty's Stationery Office.
Sutherland, N. S. & Mackintosh, J. (1964). Discrimination learning: non-additivity of cues. *Nature, Lond.*, **201**, 528–30.
Warren, J. M. & Brookshire, K. H. (1959). Stimulus generalization and discrimination learning by primates. *J. Exp. Psychol.* **58**, 348–51.
Warren, J. M. & Hall, J. F. (1956). Discrimination of visual patterns as a function of motivation and frequency of reinforcement. *J. Genet. Psychol.* **88**, 245–9.

(*Manuscript received 7 April* 1964)

Brit. J. Psychol. (1965), 56, 1, *pp.* 95–112
Printed in Great Britain

VISUAL IMAGERY PRODUCED BY RHYTHMIC PHOTIC STIMULATION: PERSONALITY CORRELATES AND PHENOMENOLOGY

By SANFORD J. FREEDMAN
Tufts University

AND PATRICIA A. MARKS
*Radcliffe College**

This study is concerned with relationships between descriptions of visual imagery produced by rhythmic photic stimulation and a number of personality tests. Individuals who manifested the ability to suspend their generalized reality-orientation described more imagery; imagination and suggestibility also seemed to be important. Subjects' expectations about what they would see influenced their reports, although comprehension of the experimental design, fatigue and motivation were not relevant. The correlated personality variables indicate a close relationship with other types of visual imagery; the phenomenology fits a synthesized description of sensory deprivation, mescaline, and hypnagogic imagery.

The photic stimulator has long been used in clinical electroencephalography. As a by-product of this work it has been found that rhythmic photic stimulation almost always produces a wide range of visual imagery. The principal object of the present study was to investigate the relationship between certain personality variables and the visual imagery obtained by rhythmic photic stimulation. Secondly, the experiment attempted to relate photic stimulation imagery to other types of imagery experience more difficult to study in the laboratory—specifically, hypnagogic imagery, sensory deprivation, hallucinations, and mescaline visions.

Previous work on these problems falls into two main categories: first, comparisons of the phenomenal qualities of the imagery occurring under different conditions; and, secondly, investigations of the personality variables related to imagery production.

(i) *Photic stimulation imagery compared with hypnagogic, sensory deprivation, and mescaline imagery*

Experimenters who have investigated photic stimulation imagery (Blum, 1956; Smythies, 1959; Walter, 1953; Walter & Walter, 1949) have described four central characteristics of the imagery: colour, movement, geometrical patterns, and meaningful images (the last reported by only very few subjects). Certain shapes such as circles, spirals, checkerboards, and waves have been mentioned with great regularity (Blum, 1956; Smythies, 1959). Smythies (1960) concluded that these characteristics of photic stimulation imagery are similar to the imagery occurring in the first stages of mescaline intoxication.

Klüver (1942) had also noted that 'some or all of the form constants found in mescaline hallucinations are also found in certain hypnagogic hallucinations' and in a rotating disk condition similar to photic flicker.

* Present address, Department of Social Relations, Harvard University.

Bexton, Heron & Scott (1954) were the first to report that sensory deprivation produced 'visual phenomena...quite similar to what have been described for mescal intoxication, and to what Grey Walter has...produced by exposure to flickering light' (p. 12). The range of phenomena they described was identical with that reported in the photic stimulation literature. Freedman, Grunebaum, Stare & Greenblatt (1962) concluded that sensory deprivation imagery is much closer to hypnagogic and hypnopompic imagery than to psychotic hallucinations.

Perhaps more convincing evidence of the similarity between the visual phenomena of hypnagogic, sensory deprivation, mescaline, and photic stimulation are the interactive and summative effects of the various conditions. For example, Smythies (1960) has found that photic stimulation potentiates the effect of mescaline and that 'hallucinations' can be induced by photic stimulation in combination with a smaller dose of mescaline than would otherwise be necessary to achieve the same effects.

Thus the literature contains evidence for phenomenological resemblances and functional relationships among hypnagogic, sensory deprivation, and mescaline imagery.

(ii) *Possible personality factors*

Smythies (1959) and Blum (1956) have reported that photic stimulation imagery patterns differ among subjects and that individual imagery patterns remain stable over long periods of time. But the photic stimulation literature is very sparse on the subject of personality factors associated with this imagery. However, since hypnagogic, sensory deprivation, mescaline, and photic stimulation imagery are, phenomenally, very much alike, it appeared that it might be fruitful to investigate the personality variables associated with imagery in these other situations in order to consider their relevance for photic stimulation imagery.

Goldberger and Holt have worked extensively with personality variables affecting the production of imagery during sensory deprivation (Holt & Goldberger, 1959, 1960; Goldberger & Holt, 1961*b*). They identified a syndrome which constituted an adaptive response to the isolation situation. This syndrome included imagery production and controlled and accepted primary process in the isolation situation. Its presence was positively correlated with an artistic, sensitive and creative self-concept. Imagery alone was found to be significantly correlated with controlled primary process during isolation. It was also highly related to vividness of imagery in everyday life in a student sample. Finally, the Rorschach colour sum measure was positively related to imagery in a group of actors, although not in the student group. McKellar (1957) pointed out that hypnagogic images and hallucinations have been used by artists as source material. Thus, sensory deprivation imagery and hypnagogic imagery have both been linked with aesthetic tendencies and creative imagination. Freedman *et al.* (1962) found a direct relationship between sensory deprivation and hypnagogic imagery. They reported a significant positive correlation between histories of hypnagogic imagery and reports of imagery from their sensory deprivation subjects. Ardis & McKellar (1956) came to a similar conclusion with respect to mescaline and hypnagogic imagery.

Camberari (1958) identified a syndrome adaptive to an isolation situation (im-

mersion in a water tank): like Goldberger and Holt's syndrome, it included a greater frequency of 'hallucinations'. The adaptive syndrome also included greater production of and tolerance for all regressive phenoma; less concern with reality orientation, i.e. a greater ability to ignore external factors that reinforce reality; and more reported affective reaction to the situation. These characteristics are analogous to what Goldberger and Holt termed 'controlled primary process' in the adaptive syndrome which they identified.

The distinguishing feature of Camberari's study was that he was able to predict differences in reaction to the isolation situation by dividing his sample into a 'suggestible' and a 'non-suggestible' group on the basis of a battery of tests. The adaptive syndrome just described was typical of the suggestible group, while the maladaptive syndrome characterized the non-suggestible group. Thus, suggestibility was positively correlated with a greater number of reports of visual imagery during isolation.

Suggestibility, as the term is used by Camberari, seems to be equated with hypnotizability. But Orne (1959) distinguished two quite different subjective components of the hypnotic state: one is a compulsion to follow cues given by the hypnotist; the other is perceived discontinuity from the normal waking state. Shor (1959) described the discontinuity from waking experience typical of the trance state by his conception of a temporary fading of the 'generalized reality orientation', defined as 'a structured frame of reference in the background of attention which supports, interprets, and gives meaning to all experiences' (p. 585). He then defined a good hypnotic subject as 'a person who has the ability to give up voluntarily his usual reality orientation to a considerable extent' (p. 598). Considering this ability as a component of Camberari's suggestibility provides an explanation of why the suggestible subjects were also less concerned with reality-orientation and more willing to ignore external factors which would reinforce reality.

A sensory deprivation study by Leiderman (1960) produced data which may possibly confirm the association between imagery and an ability to ignore external features of the environment. With six subjects, he found imagery production to be significantly correlated with field-independence as measured by the Gottschaldt test. He interpreted this result to mean that imagery is related to independence of the external environment.

In Freudian terms, Shor's 'generalized reality orientation' performs many of the functions of secondary process thinking. Similarly, the ability temporarily to give up this reality orientation corresponds to a temporary regression to primary-process functioning (Shor, 1959). Thus, it is understandable that Camberari's suggestible subjects were more productive of regressive phenomena, including 'hallucinations', and that 'controlled primary process' was associated with the occurrence of visual imagery in Goldberger and Holt's samples.

Moreover, one can also connect primary process thinking with emotional reactivity (as opposed to the dominance of logical thought processes). This interpretation could explain why a Rorschach colour response, which represents an emotional responsiveness to the environment, might also be correlated with imagery production (Holt & Goldberger, 1959; Goldberger & Holt, 1961*a*, *b*). If one takes ability to suspend the 'generalized reality orientation' as a component of hypnotizability and a factor in the production of visual imagery in the sensory deprivation situation, as

the literature suggests, then Sutcliffe's (1958) work relating hypnotizability to imagery activities is of interest. He found that hypnotizable subjects scored significantly more highly on several imagery tests, including one devoted to visual imagery exclusively.

In the light of these findings, it is interesting to see references in the literature to similarity between photic stimulation effects and the hypnotic trance state. Blum (1956) stated that a few of his eleven subjects reported trance effects from the flickering light, while Kahn (1954) found that simultaneous visual, auditory and tactile rhythmic stimulation could produce 'in a relatively short time, a hypnotic-type stupor' (p. 103). Kroger & Schneider (1959) used photic stimulation for the specific purpose of hypnotic trance induction.

Thus the concept of the 'generalized reality orientation', developed for hypnotic phenomena, seemed to explain and give unity to diverse experimental findings on the relationship between imagery and personality. Evidence concerning a similarity between the effects of photic stimulation and hypnotic trance states indicated further that this construct might furnish a useful theoretical framework for the present experiment.

In summary, the literature indicates that:

(1) There are similarities among the visual phenomena of hypnagogic states, sensory deprivation, mescaline, and photic stimulation.

(2) Photic stimulation can produce an effect similar to a hypnotic trance state.

(3) There are stable individual differences in photic stimulation imagery patterns among subjects.

(4) Sensory deprivation imagery is related to the following personality variables:

(*a*) controlled and accepted primary process; emotional responsiveness;
(*b*) vividness of imagery in everyday life;
(*c*) history of hypnagogic experience;
(*d*) artistic and imaginative self-concept;
(*e*) suggestibility, hypnotizability;
(*f*) field independence.

(5) The concept of suspension of the 'generalized reality orientation' brings these variables into a unified frame of reference, and provides a theoretical basis for extending their relevance to photic stimulation imagery.

Theory and hypotheses

The 'generalized reality-orientation' is conceived as a cognitive frame of reference which constitutes a person's ordinary 'reality'. From this most general structure 'are derived various concepts and functions, some of which are reality-testing... cognition of world,...logic,...various inhibitions, conscious fears and defenses' (Shor, 1959, p. 589). In relation to Freudian theory, these functions are 'roughly equivalent to the cognitive components of the ego or the secondary process orientation' (*ibid.*).

The fading of this generalized reality-orientation into a more distant background of awareness defines what Shor calls 'trance' (Shor, Orne & O'Connell, 1962). In Freudian terms, it refers to many of the same phenomena covered by the concept of

primary process functioning and 'regression in the service of ego'. These phenomena are viewed as having a 'trance' component in so far as they occur in relative isolation from ordinary reality. Hypnosis is one such phenomenon. Other states which Shor feels may fit into this category are hallucinogenic drug experiences, sensory deprivation, and the inspirational phase of creativity (Shor *et al.* 1962).

From the hypothesis that the hypnotic state contains a 'trance' component, 'an important deduction...is the prediction that most individuals who can readily become profound hypnotic subjects have had many profound "hypnotic-like" experiences which have occurred naturally in the normal course of living. The theory supposes that these individuals have the ability to suspend their usual generalized reality-orientation so that "hypnotic-like" experiences can occur' (Shor *et al.* 1962, p. 55). This deduction can also be extended to include other phenomena which have a 'trance' component. The present experiment sought to demonstrate that individuals with an ability to suspend their generalized reality orientations, that is, individuals who have had many naturally occurring 'hypnotic-like' experiences are more productive of photic stimulation imagery.

If visual imagery in general demands an ability to suspend one's generalized reality-orientation, then the occurrence of visual imagery experiences in a person's ordinary life should be related to his imagery production under photic stimulation. This was another hypothesis of the present study. Shor indicates that suggestibility in everyday life situations and artistic experiences (the latter a phenomenon included in 'regression in the service of the ego') also have a 'trance' component. To this extent then, they should relate to photic stimulation imagery, and such relationships constituted additional hypotheses of the present experiment.

Another hypothesis based on the assumption that imagery under photic stimulation involves the ability to suspend one's generalized reality-orientation was that people who, under ordinary circumstances, are relatively independent of their environment as a cognitive frame of reference (field-independent) are more able to let this frame of reference slip away; and, therefore, such people are likely to be more productive of photic stimulation imagery.

Shor suggests that freedom in emotional expression is related to the suspension of the generalized reality-orientation (Shor *et al.* 1962). Emotion can also be considered a primary process activity. The variable of emotional responsiveness, as an instance of controlled primary process or voluntary suspension of the generalized reality-orientation, was therefore of interest in the present experiment.

The implication throughout the preceding discussion has been that the relevant imagery dimension is quantitative rather than qualitative. One reason for the choice of quantity rather than quality was that Blum (1956) found reports of meaningful images under photic stimulation to be associated with abnormality. Since the present experiment employed a supposedly normal sample and since the personality variables used did not have a normal–abnormal dimension, the particular content or meaning of the images was not considered relevant. Moreover, previous experimental work which provided the starting point for this study used quantity rather than quality in relating imagery to other variables.

Finally, it should be made clear that the above hypotheses apply to the formal rather than the chromatic aspect of photic stimulation imagery. Both in the

literature and in common sense thinking, the minimum criterion for an image is that it should have a shape. Since the word imagery seemed to imply form as a minimum criterion, it was possible to formulate hypotheses about this aspect of the imagery on the basis of previous experimental work. This was not the case with the colour aspect of the imagery. Although no predictions could therefore be made, the relationship between colour reports under photic stimulation and the personality variables was also investigated.

EXPERIMENTAL DESIGN

Each subject was tested once over a range of flicker frequencies to obtain a sample of the formal and chromatic aspects of his imagery. The results of a pilot study had indicated that one series of exposures to the flashing light over an adequate range of frequencies would be sufficient to obtain a representative sample of the variety and amount of imagery for a given subject. After this series of exposures, a form suggestion procedure was used in order to test hypotheses about suggestibility.

One problem of design was that if imagery was responsive to explicit suggestions, as had been indicated in the pilot study, it seemed likely that it would also be affected by the implicit clues offered by the experimental procedure itself—what Orne has termed the 'demand characteristics' of the experiment. Thus, the degree to which suggestibility affected imagery reports might depend upon the way in which the subject perceived the experimental situation. Consequently, it seemed necessary to assess and control for 'demand characteristics'. The strategy used for this purpose was Orne's (1959): subjects were intensively questioned concerning their expectations before the experiment, and their ideas about the purpose of the photic stimulation procedure after it was over. It was assumed that subjects who perceived the experiment in a certain way would manifest similar experimental behaviour if 'demand characteristics' were a determining factor.

Another methodological problem was that imagery reports might be influenced by verbal ability as much as, or more than, by the visual experience itself. The influence of verbal ability was assessed by asking subjects how adequately they felt they had been able to describe their visual experience. It was assumed that if this rating was positively correlated with the amount of imagery reported, verbal ability had contributed significantly to the supposedly visual results.

METHOD AND PROCEDURE

(i) *Subjects*

Twenty paid subjects, ten male and ten female, were recruited through the student employment offices of three nearby colleges. Ages ranged from 18 to 25, with a mean of 20·3 years.

(ii) *Photic stimulation sessions*

The experiment involved two sessions one or two days apart. At the start of the first session the subjects were told only that the experiment would involve an electroencephalogram, a flashing light, and a description of what they saw while the light was on. Before photic stimulation began, subjects were questioned about their motivation for participating and their expectations concerning the experiment. They were then given a clinical EEG by a hospital technician in order to screen out those who exhibited adverse reactions to the light. The clinical EEG and the experiment took place in a dark room with the subject recumbent on a bed, eyes closed. After the clinical EEG, the subject was instructed that while the light was on he was to describe everything he saw as carefully and accurately as possible. He was also told to look straight ahead with his eyes closed, not to move, and to relax as much as possible. A Grass PS-2 Stimulator was placed about 1 in. from the subject's nose so as to cover the visual field; relative intensity was set at 16. In this part of the photic stimulation procedure, called the 'spontaneous reporting condition', the duration of each exposure to the light was 10 sec. The first exposure was at a frequency of 1 f.p.s. (flash per second) with later exposures progressing successively up to a frequency of 20 f.p.s. The amount of time between exposures was determined

primarily by the length of the subject's description, which was allowed to continue after the light went off, if necessary.

In the second session of the photic stimulation procedure, the 'suggestion condition', subjects again described everything they saw, but in addition were asked to watch particularly for a specific form on successive 30 sec. exposures. The suggested forms were: (1) circles, (2) squares, (3) spirals, (4) stars, (5) flowers, (6) animals and (7) people. The flicker frequency was 15 f.p.s. When a subject reported having seen a suggested form, he was asked whether it had been in motion, whether he had ever seen it before, and how he had described it (if he had seen it previously). The subject was asked to rate himself on a 5-point scale for fatigue, at the beginning and the end, and in the middle of the photic stimulation procedure.

After the suggestion procedure, the experimenter asked the subject a series of questions (the photic stimulation inquiry). These questions were concerned with the phenomenal qualities of the imagery, the subject's affective reaction to the experience, and his conception of the nature of the experiment.

(iii) *Coding the imagery*

Visual imagery reports during the 'spontaneous' condition were coded for form and colour. A form response was defined as any mention of shape which implied at least two dimensions, that is, dots and lines were excluded. This definition included objects of the real world (e.g. rocket, pinwheel) as well as geometric figures (e.g., square, circle). The imagery scores were:

(*a*) *Variety of form*—the number of *different* form responses given over all twenty of the 10 sec. trials in the spontaneous reporting condition. For example, the response 'circle' would be counted only once even if it were repeated several times. Actual form scores ranged from 1 to 19.

(*b*) *Variety of objects*—the number of different forms reported over the same twenty trials which referred to objects of the external world; i.e. geometric figures were excluded.

(*c*) *Frequency of colour*—the *total number* of colour responses given over all twenty of the 10 sec. exposures. If a given colour response was repeated on the same trial it was counted only once. Actual colour scores ranged from 0 to 45.

Imagery during the 'suggestion condition' was scored for positive responses to suggestion. A positive response was defined as any mention of the suggested form during the following exposure. The following measures were used:

(*a*) *Form suggestibility*—the sum of all positive responses to suggestion. Each form was assigned a weight in accordance with the empirically determined 'difficulty' of seeing it.

(*b*) *New forms under suggestion*. This measure was based on the same scoring system. However, scores for those suggested forms which subjects reported having seen before the suggestion condition were subtracted.

(iv) *Testing session*

All personality tests were individually administered by a single experimenter. All subjects were given the following tests in the order listed:

(*a*) *A modification of the Gottschaldt embedded figures test*—(Witkin, 1950)

This test, which involves extracting a simple geometric figure from a complex context of lines and colour, was employed as a measure of field-independence to test the hypothesis that this variable is positively related to amount of photic stimulation imagery. According to Gardner *et al.* (1959), one group of figures (cluster I) sets a more purely perceptual task than the other group of figures (cluster II), which demands conceptual activity as well. Consequently, scores were computed separately for each cluster. Scores consisted of the total amount of time taken to find all the embedded figures.

(*b*) *Memory for positions test*

This test, one of those used by Camberari (1958), was employed as a measure of suggestibility to test the hypothesis that visual imagery is positively related to suggestibility. It involved remembering how many letters on one card were in the same position as on another card seen immediately before. The correct answer in each case was none. Suggestibility was measured by the number of letters guessed.

(c) Rorschach test

Three Rorschach cards (II, IX, X), containing varying amounts of chromatic colour, were presented to each subject. The protocols were scored by an experienced clinical psychologist and a colour sum ratio was computed according to the formula $(0{\cdot}5FC + CF + 1{\cdot}5C)/\text{total } R$. This score was taken as a measure of the subject's ability to respond spontaneously to the environment in emotional terms and was used to test whether or not more emotionally responsive subjects would see more imagery. (Rorschach results are based on nineteen out of the twenty subjects.)

(d) Barron–Welsh art scale (Welsh, 1959)

This scale was employed as an objective measure of artistic ability, sensitivity and creativity (Barron, 1953; Welsh, 1959) to test the hypothesis that these traits are associated with high imagery production. The scale consists of sixty-two out of the 400 black and white drawings contained in the Welsh Figure Preference Test. Subjects rate each drawing 'like' or 'dislike'. The scale was scored according to the key provided by Welsh (1959).

(e) Shor 'Hypnotic-like' experiences questionnaire (Shor, 1959)

This questionnaire was conceived as a measure of 'the ability to lose temporarily or to suspend voluntarily the relative functioning of the generalized reality-orientation' (Shor, 1960, p. 162). Scores have been found to correlate with hypnotizability, and it was hypothesized that they would also correlate with the amount of photic stimulation imagery. In general, the questions ask whether the subject has ever had various mystical-type experiences not induced by special conditions such as drugs or hypnosis; for example, 'Have you ever been able to quiet down your mind, construct a new imaginary world and feel for the time that it was real?' In accordance with Shor's methods, scores consisted of the number of hypnotic-like experiences the subject reported having had at least once.

(f) Imagery and creativity questionnaire

This questionnaire includes nineteen items about past imagery experience (hypnagogic imagery and dreams), artistic inclinations, self-ratings of imaginative and creative abilities, and empathy experience. On the basis of interviews conducted during the pilot study, these items seemed most promising as predictors of photic stimulation imagery.

Subjects responded to each item on a three-point rating scale. An overall questionnaire score was based on eleven of the nineteen times. The selected items are all concerned with visual imagery and creativity, and they constitute a scale referred to as 'imagery and creativity'. In addition, each of the nineteen items was considered separately in an item-by-item analysis.

RESULTS

(i) *The relation between personality variables and photic simulation imagery*

The main body of results relating photic stimulation imagery to the personality variables is presented in Table 1.

(a) Interrelation among imagery variables

The colour and form aspects of imagery were independent, i.e. the correlations between form and colour responses did not approach statistical significance.

Form suggestibility correlated significantly with the other form measures. Although form measures and form suggestibility are based on different observations, it was possible that people saw the same things under suggestion that they had

already seen in the 'spontaneous' condition. However, the high correlation ($r_s = 0{\cdot}86$) of form suggestibility with the number of new forms (i.e. forms reported as not seen in the spontaneous condition) shows that this relationship between forms reported spontaneously and those reported under suggestion is not based on a simple repetition of imagery from trial to trial (Table 2).

Table 1. *Matrix of rank order correlations: photic stimulation imagery and personality variables* ($N = 20$)

	2	3	4	5	6	7	8	9	10
1. Variety of form	0·56**	−0·08	0·71**	0·47*	0·39*	0·60**	−0·07	0·21	0·29
2. Variety of objects	—	0·05	0·45*	0·36	0·30	0·43*	−0·11	0·00	0·51*
3. Frequency of colour	—	—	0·12	0·22	−0·22	0·18	−0·19	0·63**	0·33
4. Form suggestibility	—	—	—	0·41*	0·45*	0·38	−0·25	0·37	0·21
5. Imagery and creativity	—	—	—	—	0·18	0·35	0·05	0·12	0·03
6. 'Hypnotic-like' experiences	—	—	—	—	—	0·15	0·01	0·22	0·12
7. Rorschach colour sum ratio ($N = 19$)	—	—	—	—	—	—	−0·05	0·03	0·03
8. Barron–Welsh art scale	—	—	—	—	—	—	—	−0·11	0·08
9. Embedded figures test, cluster I	—	—	—	—	—	—	—	—	0·38*
10. Embedded figures test, cluster II	—	—	—	—	—	—	—	—	—

* $P < 0{\cdot}05$; ** $P < 0{\cdot}01$ (one-tailed tests).

Table 2. *Rank order correlations: variety of form responses, form suggestibility, new forms under suggestion* ($N = 20$)

	2	3
1. Variety of form	0·71**	0·59**
2. Form suggestibility	—	0·86**
3. New forms under suggestion	—	—

** $P < 0{\cdot}01$ (one-tailed tests).

(*b*) *Interrelation among personality variables*

From the lower right-hand segment of Table 1 it can be seen that there were no significant correlations among the personality measures (except between the two parts of the Gottschaldt test).

(c) *'Hypnotic-like' experiences*

A demonstrated ability to suspend temporarily one's 'generalized reality-orientation', as manifested in frequency of 'hypnotic-like' experiences reported in the Shor questionnaire, was significantly correlated with variety of form and form suggestibility. Thus, the hypothesis that individuals who have a general ability to suspend their generalized reality-orientation are more productive of photic stimulation imagery was confirmed with respect to the form aspect of the imagery. The fact that frequency of 'hypnotic-like' experiences correlated somewhat more strongly with form suggestibility than with form or colour is an indication that 'hypnotic-like' experiences are more closely related to phenomena occurring in a situation more like that of hypnosis. Assuming that a 'trance' component (as defined by Shor) is common to both hypnosis and imagery, this higher correlation

between 'hypnotic-like' experiences and form suggestibility suggests an additional component, suggestibility, which is called into play in both hypnosis and form suggestibility, but not in spontaneous reporting of imagery.

(*d*) *Imagery and creativity*

Self-ratings on the eleven imagery and creativity items selected from the questionnaire correlated significantly with variety of form and form suggestibility. Partial correlation, however, indicated that a large part of the correlation between 'imagery and creativity' and form suggestibility, could be attributed to a relationship between variety of form and form suggestibility. This result suggests that the 'imagery and creativity' items pertain to special imagery abilities distinct from 'trance ability' (although perhaps occurring in a trance state). These abilities would be more relevant to 'spontaneous' imagery than to imagery suggestibility. Artistic sensitivity and creativity, as measured by the Barron–Welsh Art Scale, was not significantly correlated with any of the photic stimulation measures, however.

(*e*) *Field-independence*

Field-independence, cluster I (the perceptual group of embedded figures) was significantly correlated with frequency of colour. Field independence, cluster II (the conceptual group of embedded figures) was significantly correlated with variety of objects.

The hypothesis that field-independent people see more imagery because they more readily abandon their generalized reality-orientation was not confirmed by the results if variety of form is taken as the criterion of imagery production. However, perceptual field-independence (cluster I) correlates significantly with the colour measures. There does not seem to be an obvious explanation for this result since attention to colour is detrimental to performance on the Embedded Figures Test.

(*f*) *Emotional responsiveness*

The Rorschach colour sum ratio, conceived as a measure of emotional responsiveness to external stimuli, was positively correlated with all the measures of photic stimulation imagery. The correlations of the colour sum ratio with variety of form and variety of objects were statistically significant.

The strong positive relationship between Rorschach colour sum ratio and photic stimulation imagery supports the hypothesis that emotional responsiveness to external stimuli is a factor in imagery experience. On the assumption that emotion is a primary process function, this result also supports the theoretical relationship between the suspension of the generalized reality-orientation and 'regression in the service of the ego', as well as giving evidence that such a process is linked with imagery production. However, partial correlation indicated that most of the relationship between variety of objects and the Rorschach seems to have been a result of the relationship between variety of objects and variety of form, for when variety of form was held constant by a partial correlation, the correlation between the Rorschach and variety of objects sank to 0·23. When variety of objects was held constant, on the other hand, the correlation between the Rorschach colour sum ratio and variety of form increased to 0·80.

(*g*) *Suggestibility*

Although *t*-tests performed on the memory for positions scores did not yield statistical significance, the differences between the means of high and low suggestibility subjects were consistently in the predicted direction; that is, highly suggestible people tended to report a greater variety of form, a higher frequency of colour and tended to respond to form suggestion to a greater extent. These results are in accord with the finding that 'hypnotic-like' experiences are related to imagery production, and they support, to some extent, the hypothesis relating imagery and suggestibility.

(ii) *Imagery and creativity questionnaire*

(*a*) *Questionnaire and variety of form*

An item-by-item analysis of all nineteen of the questionnaire items showed that two of the items were able to predict variety of form scores significantly and were therefore presumably responsible for most of the correlation computed on the basis of the total score of eleven items. The first item in Table 3 shows that subjects who considered themselves quite imaginative, as opposed to those who thought of themselves as only a little imaginative, produced a significantly greater variety of imagery.

From the second item in Table 3, it can be seen that those who reported a previous

Table 3. *Average variety of form and frequency of colour in photic stimulation imagery as a function of self-ratings on imagery and creativity questionnaire*

Do you consider yourself to be an imaginative person?

	(*a*) Yes, quite ($N = 8$)	(*b*) Yes, a little ($N = 12$)	(*c*) No ($N = 0$)	F	P
Form	10·75	5·33	0	2·58	< 0·01
Colour	16·38	25·17	0	0·14	N.S.

Many people see quite vivid images when falling asleep or waking up, usually (but not always) with their eyes closed. Has this happened to you?

	(*a*) No ($N = 8$)	(*b*) Yes, once or twice ($N = 3$)	(*c*) Yes, more frequently ($N = 9$)	F	P
Form	3·88	11·67	9·33	4·75	< 0·05
Colour	24·75	19·67	19·89	—	N.S.

How frequently do you paint or draw?

	(*a*) Often ($N = 4$)	(*b*) Infrequently ($N = 4$)	(*c*) Never ($N = 12$)	F	P
Form	7·00	7·25	7·75	—	N.S.
Colour	34·25	20·25	17·83	22·43	< 0·005

Do you consider yourself to be a creative thinker?

	(*a*) Yes, quite ($N = 4$)	(*b*) Yes, somewhat ($N = 11$)	(*c*) No ($N = 5$)	F	P
Form	9·25 4·7	6·73	7·80	—	N.S.
Colour	11·25 1·9	23·27	26·40	9·42	< 0·005

history of hypnagogic and hypnopompic imagery differed in variety of photic stimulation imagery. The discriminating factor seems to be whether or not subjects reported ever having had previous imagery. Subjects who reported never having had hypnagogic or hypnopompic imagery differed significantly from subjects who reported 'once or twice' and from those who reported having had it 'more frequently'. The difference between the means of 'once or twice' and 'more frequently' was not significant.

(b) Questionnaire and frequency of colour

Several items of the questionnaire significantly differentiated among subjects with regard to amount of colour in their photic stimulation imagery. These items are all different from those which predicted the form aspect of imagery, although their general nature is similar. It can be seen from the third item in Table 3 that subjects who painted or drew most frequently reported the greatest amounts of colour in their imagery ($P < 0{\cdot}005$). The results of the fourth item in Table 3 indicate that self-ratings of creativity were also positively associated with frequency of colour in photic stimulation imagery.

(iii) *Situational and control variables*

The following situational and 'control' variables had no effect on imagery production: self-ratings of fatigue, perceived ability to verbalize the visual experience, perceived control over the imagery, and motivation for participating in the experiment. Motivation to see the desired forms, like motivation for the experiment, did not affect form suggestibility.

Table 4. *Average frequency of variety of form, frequency of colour, and form suggestibility as a function of whether subjects perceived or failed to perceive imagination as an experimental variable*

	Named imagination as variable			
	Yes ($N = 5$)	No ($N = 15$)	t	P
Variety of form	8·60	7·13	—	N.S.
Frequency of colour	31·80	18·90	1·87	$< 0{\cdot}05$
Form suggestibility	7·14	6·33	—	N.S.

In addition, the 'demand characteristics of the experimental situation' did not appear to have a decisive influence on the results. Subjects succeeded in perceiving two variables which were, in fact, a concern of the experimenter: imagination and suggestibility.

This perception, however, did not significantly increase experimental imagery scores for variety of form and suggestibility (Table 4). Moreover, perception of imagination as a variable was not significantly associated with a higher self-rating of imaginativeness on the questionnaire. Colour, on the other hand, did seem to have been affected by the perception of imagination as an experimental variable. These results indicate that the present design was such as to make the effect of situational determinants negligible, thus emphasizing the personality variables. Specific

expectations, on the other hand, seemed to exert a very strong influence on imagery (Table 5).

The one subject who expected to see 'things' under the light had the highest variety of form score, deviating from the mean by 12·11 ($P < 0·005$). The three subjects who expected to see colours under the light had mean colour scores significantly greater than those of subjects whose expectations did not refer to colour (Table 5).

Table 5. *Average frequency of form and colour in imagery of people with different expectations before experiment*

	Expectations			
	Expected to see 'things' ($N = 1$)	No expectations related to seeing 'things' ($N = 19$)	Z	P*
Variety of form	19·00	6·89	2·63	< 0·005
	Expected to see colours ($N = 3$)	No expectations related to colour ($N = 17$)	t	P
Frequency of colour	39·33	18·24	2·67	< 0·01

* One-tailed, for normal deviate.

Table 6. *Average frequency of imagery for subjects with different expectations about suggestion procedure*

	Expected to see suggested forms				
	Yes ($N = 6$)	Some of the forms ($N = 7$)	No ($N = 7$)	F	P
Variety of form	10·50	8·14	4·27	2·87	N.S.
Form suggestibility	10·83	7·00	3·00	6·52	< 0·01

Expectations about seeing suggested forms, as assessed at the end of the photic stimulation procedure, seem to be a strong factor in form suggestibility. The difference between subjects who reported expecting to see the suggested forms and those who did not report such expectations is significant at the 0·01 level (Table 6). However, interpretation of these results must be tempered by the fact that this inquiry about expectation was made after the suggestion trials had occurred. Consequently, remembered expectations may have been affected by subjects' intervening experience. If, however, expectation did have an effect, it seems that such expectations were based in part on past form experience, for subjects who expected to see the suggested forms had seen a greater variety of form in the first place (Table 6), although this difference did not attain statistical significance.

(iv) *Phenomenology*

(*a*) *Affective reactions*

Subjects' affective reactions to the photic stimulation could be classified as positive, negative or neutral. Although affective reactions did not differentiate subjects

to a significant degree with regard to amount of form and colour in their imagery, it is at least interesting that the two different questions 'Did you enjoy your experience under the light?' and 'Did you find the light at all annoying?' produced results related in the same way with respect to form. That is, subjects who were most positive in their reactions reported the most form, the most negative reported the next greatest amount, and the most neutral reported the least. This result lends credibility to the link between emotional responsiveness (as measured with the Rorschach) and imagery, for subjects who had either positive or negative affective reactions to the experience reported a greater variety of form than those whose reactions were neutral.

(b) *Perceived relationship with past experience*

Subjects viewed the photic stimulation phenomena as related to a variety of past experiences. Eight subjects associated it with various previous exposures to lights. Four subjects related the photic stimulation experiment to drug experiences of various sorts (peyote, marijuana, mescaline, nitrous oxide). Two subjects mentioned visual experiences produced by closing their eyes tightly or pressing their eyes. One subject related his photic stimulation experience to hypnagogic images. Three subjects did not see any relationship between photic stimulation imagery and previous experience. These three subjects did not differ significantly from the rest of the sample in the amount of form or colour contained in their photic stimulation imagery.

(c) *Qualitative description of the imagery*

When asked to compare photic stimulation imagery with perceptions of ordinary 'reality', eleven of the twenty subjects said that none of the images was as realistic as things seen in ordinary life. Only one reported that all the images were as real as life. The most frequent reason given for knowing that the images were *not* real was their blurredness or indistinctness (seven subjects). Nevertheless, seventeen of the twenty subjects felt as if they were really seeing the photic stimulation images (as opposed to imagining them). Imagery content ranged from lines and dots to geometric figures to objects of the real world to integrated scenes (only one subject).

At the end of the photic stimulation session, subjects were asked to imagine specific forms and to compare these images with the remembered photic stimulation images of the same forms. Of the total number of comparisons, 65 % of the imagined forms seemed less real than their photic stimulation counterparts; 88 % seemed less vivid; and 77 % of the images were not 'seen' in the same way as the photic stimulation images, that is, they seemed more 'thought' than 'seen'. Thus, while photic stimulation images were 'seen' rather than imagined, they lay somewhere between imagination and open-eyed perception in their 'reality' qualities. In general, they were often less distinct than reality, although the colours were frequently more vivid or picture-like than those seen in the everyday world. Here is evidence in favour of the assumption that photic stimulation demands suspension of one's usual reality-orientation, for the reporting of an image seemed to signify that subjects were actually admitting that they saw things which at the same time they knew were not 'real'. Such a contradiction would not be tolerated by the more logical reality-testing typical of secondary process function, one aspect of the generalized reality-orientation.

In general, the phenomenal qualities of the imagery fit a synthesized description of sensory deprivation, mescaline and hypnagogic imagery offered by Freedman *et al.* (1962). However, the description is somewhat too simple to fit all the experience of all the photic stimulation subjects. The way in which photic stimulation imagery compared with these other types of imagery on five main points common to them all is presented below.

(1) *Autonomy.* Half the subjects felt they had some sort of control over the imagery, but this control was not perceived as complete. None mentioned being able to control occurrence of the images, as opposed to their content. On the whole, it seems that photic stimulation images appeared at least partially autonomous to all subjects.

(2) *Irrelevance to thoughts of subjects.* This quality could be inferred from the relatively small number of meaningful (i.e. non-geometric) images which were reported.

(3) *Originality.* Originality in the sense of unreproductiveness was a predominant quality of much of the imagery. Only one of the subjects felt that the images 'reproduced' the reality of the external world.

(4) *Unnaturally vivid colour.* Three subjects reported that the images were coloured more vividly than things seen in ordinary life. Nine other subjects perceived at least some of the colours as extraordinary, but in what way was uncertain. Only one subject thought all the colours were those of the real world. Thus, it can be concluded that the colours were on the whole 'unnatural', although whether always in the direction of extraordinary vividness is uncertain.

(5) *Brief duration.* This was an outstanding quality of the imagery as manifest in the inability of twelve subjects to describe accurately the imagery because of rapid changes.

Thus four of the five outstanding characteristics which have been reported for sensory deprivation, mescaline, and hypnagogic imagery held true for most of the images for most of the subjects. The fifth characteristic, unnatural vividness of colour, was typical of most of the subjects with regard to the unnatural aspect. In addition, ordinary vividness was noted by several subjects.

Discussion

On the whole, the results seem to confirm the usefulness of suspension of the generalized reality-orientation (or 'trance') as a theoretical framework within which to discuss the imagery produced by rhythmic photic stimulation. The relationships obtained using variety of form as the principal imagery measure indicated the importance of three clusters of variables which can be derived from this framework: 'hypnotic-like' experiences in ordinary life; imagination and past imagery experience (notably hypnagogic and hypnopompic imagery); and emotional responsiveness to external stimuli. The place of imagination is not so clear as the others; the Barron–Welsh Art Scale, which should theoretically have been related to self-rating of imaginative ability and hence to photic stimulation imagery, was not so related.

The colour aspect of photic stimulation imagery did not fit into such an orderly pattern. Except for some indication that colour may be related to creative thought

and artistic ability, the relationships found in this study do not have much order. The pilot study had indicated that colour reports were very much influenced by suggestion. One possible interpretation of the result regarding colour is that there is a kaleidoscope of colours generally present while the light is flickering and the particular colour attended to at a given moment can be a matter of a multitude of factors unless a specific set is induced. Perhaps such a set was induced by the 'demand characteristics' of the present experiment, for those subjects who perceived imagination as an experimental variable reported significantly more colour in their imagery than subjects who did not have such a perception.

The relationship between hypnotizability and imagery extends the generality of the reports of McBain (1954), Shor (personal communication, 1962), and Sutcliffe (1958) that waking imagery is related to hypnotizability. This relationship between 'hypnotic-like' experiences and imagery gains credibility from the observations of 'hypnotic-like' trances produced by rhythmic photic stimulation, sometimes in combination with rhythmic stimulation in other modalities (Blum, 1956; Kahn, 1954; Kroger & Schneider, 1959). This result also agrees with Camberari's (1958) finding that suggestible subjects have more visual experience in a sensory deprivation situation.

The relationship found between self-rating of imaginative ability and photic stimulation imagery agrees in a general way with Goldberger & Holt's (1961*b*) finding that an artistic, sensitive, and creative self-concept is related to a syndrome that includes imagery. However, self-ratings of artistic ability and creativity in particular did not relate significantly to imagery in the present study. Like that of Goldberger & Holt, the present experiment was not very successful in relating imagery to creativity on the Barron–Welsh Art Scale. In general, it appears that self-ratings of imaginative ability correlate with imagery whereas objective tests do not. The most parsimonious explanation is that self-ratings and objective tests of these variables have little to do with each other. This interpretation is supported by the lack of correlation between the imagery and creativity items of the questionnaire and the Barron–Welsh Art Scale.

The link between hypnagogic imagery experience and photic stimulation imagery supports the general notion of a relationship between the imagery occurring in various situations (Ardis & McKellar, 1956; Bexton *et al.* 1954; Freedman *et al.* 1962; Galton, 1883; Klüver, 1942; Smythies, 1960). In particular, it extends the generality of the finding of Freedman *et al.* (1962) that previous hypnagogic imagery experience is a determinant of imagery in a sensory deprivation situation.

The relationship found between emotional responsiveness to external stimuli and imagery supports Goldberger & Holt's (1961*b*) finding of a relationship between colour responses on the Rorschach and imagery in their actor sample. The relation between imagery and emotion also supports in a more general way their finding that 'controlled primary process' in the isolation situation is correlated with imagery production. Camberari (1958) also found imagery in sensory deprivation associated with other primary process or 'regressive' phenomena.

Finally, the positive relation between field-independence and imagery obtained by Leiderman (1962) in a sensory deprivation experiment was not supported in this photic stimulation situation. Two possible explanations of this discrepancy are readily

apparent. One is that Leiderman's results, based on a very small sample of subjects may not represent a true relationship. The other is that the relationship is specific to conditions present in the sensory deprivation situation and absent in photic stimulation.

The success of the present study in relating photic stimulation imagery to personality variables which have been found relevant to imagery in other situations indicates the potential value of rhythmic photic stimulation as a laboratory method for studying phenomena which are similar to those found in areas of more practical interest. The fact that photic stimulation, unlike sensory deprivation or mescaline, produces visual imagery unaccompanied for the most part by complicated psychological and physiological reactions has implications for research whose primary interest is 'unreal' visual phenomena.

Moreover, the results of the present study have implications for the study of personality, as well as of imaginal processes. They can be viewed as belonging to the more general area of 'personality and perception' as represented, for example, by Blake & Ramsey (1951) and Witkin *et al.* (1954). The syndrome of abilities or traits identified in the present experiment is concerned with an extremely basic aspect of personality functioning, the question of what a person can accept as reality. As such, this syndrome may have relevance to a wide variety of situations which are conducive to 'unreal' experiences. For example, Klein, Gardner & Schlesinger (1962) have recently described a 'cognitive control' named 'tolerance for unrealistic experiences', which they have shown to account for individual differences in approach to unconfirmable or unrealistic stimulus material presented in situations such as apparent movement, distortion of the visual field by aniseikonic lenses and the Rorschach test. This 'tolerance for unrealistic experiences' is extremely similar to and should apply to the same phenomena as Shor's (1959) construct, the generalized reality-orientation.

If this is the case, then the personality syndrome identified in this photic stimulation experiment—imaginativeness, emotional responsiveness, and a tendency to mystical-type experiences—may have relevance for a wide variety of situations which are conducive to 'unreal' experience, as well as for imagery of all kinds.

The authors wish to thank the following people for assistance at various stages: Mary Finn, R.N., Anna Girouard, Rudolf Kälin, Dr R. E. Shor, and Dr Justin Weiss. The work was supported by Grant AF-AFOSR 62-11 from the U.S. Air Force Office of Scientific Research.

References

Ardis, J. A. & McKellar, P. (1956). Hypnagogic imagery and mescaline. *J. Ment. Sci.* **102**, 22–9.

Barron, F. (1953). Complexity—simplicity as a personality variable. *J. Abnorm. Soc. Psychol.* **48**, 163–72.

Bexton, W. H., Heron, W. & Scott, T. H. (1954). Effects of decreased variation in the sensory environment. *Canad. J. Psychol.* **8**, 70–6.

Blake, R. R. & Ramsey, G. V. (eds.) (1951). *Perception: An Approach to Personality.* New York: Ronald.

Blum, R. H. (1956). Photic stimulation, imagery, and alpha rhythm. *J. Ment. Sci.* **102**, 160–7.

Camberari, J. D. (1958). The effects of sensory isolation on suggestible and non-suggestible psychology graduate students. Unpublished doctoral dissertation, University of Utah.

Freedman, S. J., Grunebaum, H. U., Stare, F. A. & Greenblatt, M. (1962). Imagery in sensory deprivation. In L. J. West (ed.), *Hallucinations*, pp. 108–17. New York: Grune and Stratton.

Galton, F. (1883). *Inquiries into Human Faculty and its Development*. New York: MacMillan.

Gardner, R., Holzman, P. S., Klein, G. S., Linton, H. & Spence, D. P. (1959). Cognitive control. A study of individual consistencies in cognitive behavior. *Psychol. Issues*, **1**. No. 4.

Goldberger, L. & Holt, R. R. (1961*a*). Experimental interference with reality contact: individual differences. In Solomon, P., Kubanzky, P. et. al. (eds.). *Sensory Deprivation*. Cambridge, Mass: Harvard U.P.

Goldberger, L. & Holt, R. R. (1961*b*). A comparison of isolation effects and their personality correlates in two divergent samples. Final report WADD Contract No. AF 33(616)-6103. New York Univ. Res. Ctr. for Mental Health.

Holt, R. R. & Goldberger, L. (1959). Personological correlates of reactions to perceptual isolation. WADC Technical Report 59-735. Wright-Patterson Air Force Base, Ohio.

Holt, R. R. & Goldberger, L. (1960). Research on the effects of isolation on cognitive functioning. WADD Technical Report 60-260. Wright-Patterson Air Force Base, Ohio.

Kahn, T. C. (1954). Theoretical foundations of audio-visual-tactile rhythmic induction therapy experiments. *Science*, **120**, 103–4.

Klein, G. S., Gardner, R. W. & Schlesinger, N. J. (1962). Tolerance for unrealistic experiences: a study of the generality of a cognitive control. *Brit. J. Psychol.* **53**, 41–55.

Klüver, H. (1942). *Mechanisms of Hallucinations*. In McNemar, Q. & Merrill, M. A. (eds.), *Studies in Personality*. New York: McGraw-Hill.

Kroger, W. S. & Schneider, S. A. (1959). An electronic aid for hypnotic induction. *Int. J. Clin. Exp. Hypnosis*, **7**, 93–8 (abstract).

Leiderman, P. H. (1962). Imagery and sensory deprivation: an experimental study. Tech. Rpt. MRL-TDR-62-28, Wright-Patterson Air Force Base, Ohio.

McBain, W. N. (1954). Imagery and suggestibility: a test of the Arnold hypothesis. *J. Abnorm. Soc. Psychol.* **49**, 36–44.

McKellar, P. (1957). *Imagination and Thinking*. New York: Basic Books.

Orne, M. T. (1959). The nature of hypnosis: artifact and essence. *J. Abnorm. Soc. Psychol.* **58**, 277–99.

Shor, R. E. (1959). Hypnosis and the concept of the generalized reality-orientation. *Amer. J. Psychother.* **13**, 582–602.

Shor, R. E. (1960). The frequency of naturally occurring hypnotic-like experiences in the normal college population. *Int. J. Clin. Exp. Hypnosis*, **8**, 151–63.

Shor, R. E., Orne, M. T. & O'Connell, D. N. (1962). Validation and cross-validation of a scale of self-reported personal experiences which predicts hypnotizability. *J. Psychol.* **53**, 55–75.

Smythies, J. R. (1959). The stroboscopic patterns. Part II: the phenomenology of the bright phase and after-images. *Brit. J. Psychol.* **50**, 305–24.

Smythies, J. R. (1960). The stroboscopic patterns. Part III: further experiments and discussion. *Brit. J. Psychol.* **51**, 247–55.

Sutcliffe, J. T. (1958). Hypnotic behaviour: fantasy or simulation? Unpublished doctoral dissertation, University of Sydney.

Walter, W. G. (1953). *The Living Brain*. New York: Norton.

Walter, V. J. & Walter, W. G. (1949). The central effects of rhythmic sensory stimulation. *EEG. Clin. Neurophysiol.* **1**, 57–86.

Welsh, G. S. (1959). *Preliminary Manual, Welsh Figure Preference Test*. Palo Alto: Consulting Psychologists Press.

Witkin, H. A. (1950). Individual differences in ease of perception of embedded figures. *J. Personality*, **19**, 1–15.

Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B. & Wapner, S. (1954). *Personality Through Perception*. New York: Harper.

(*Manuscript received 21 November 1963*)

Brit. J. Psychol. (1965), 56, 1, pp. 113–119
Printed in Great Britain

PUBLICATIONS REVIEWED

Two Interdisciplinary Studies for Psychologists

Human Senses and Perception. By G. M. WYBURN, R. W. PICKFORD and R. J. HIRST. Edinburgh and London: Oliver and Boyd. 1964. Pp. xii + 340. 45*s*.

Behaviourism and Phenomenology. Edited by T. W. WANN. London: University of Chicago Press. 1964. Pp. xi + 190. 37*s*. 7*d*.

Different as these two books are, they are both outstanding examples of the progress of 'interdisciplinary collaboration' in contemporary psychology.

Early in the century the main course in psychology at Cambridge was conducted by G. E. Moore lecturing to a syllabus based on Ward's famous and then still influential article in the *Encyclopaedia Britannica*. The first sentence in this syllabus was: 'The standpoint of psychology is individualistic.' Moore spent the whole of the first term (the class met three days a week) trying to discover what Ward could possibly have meant by this statement. The attempt was unsuccessful. In a discussion period at the end of the term an intelligent, but innocent, Indian student piped up: 'Mr Moore, have you ever asked Professor Ward what he meant?' The question was not unnatural since Moore and Ward lectured in adjoining classrooms and shared a common robing room. 'No,' said Moore, tersely, with an emphasis which implied that to have done so would have been an act of gross academic impropriety.

Since those days there have been advances in interdisciplinary co-operation, notably in research, considerably in the organization of conferences, but very little in undergraduate teaching.

Human Senses and Perception is accordingly a landmark. It exemplifies the way in which teachers of physiology, of psychology and of philosophy can work together to provide the essential content of a course of study of Sense Perception. Its authors are G. M. Wyburn, Regius Professor of Anatomy in the University of Glasgow (who is also the editor of the volume), R. W. Pickford, Professor of Psychology at Glasgow, and R. J. Hirst, the Professor of Logic in the same University. In Part I Prof. Wyburn gives an admirably lucid factual account of the anatomy and physiology of the senses. In Part II Prof. Pickford covers the distinctively psychological topics—the perception of space, time, movement and causality, the learning process in perception, and individual and social factors influencing perception. He concludes with an admirable, though rather too brief, summary of General Theories. An attractive feature of Pickford's section is that it covers not only the fields of more ordinary text-books, but also more out-of-the-way topics in which students are naturally and quite properly interested, such as those of psychological aesthetics and the alleged facts of paranormal perception. In Part III Prof. Hirst takes up the problems of perception which are of special interest to philosophers and which cannot be altogether evaded by students of psychology. He introduces the student to the main classical and contemporary issues: the representative theory of perception, 'sense datum' theories and common-sense theories, the analysis of perceptual consciousness and the traditional theories of body and mind with which the special problems are inextricably entangled. It is, perhaps, a pity that the publishers were allowed to describe Prof. Hirst's own contribution as unorthodox and as opposed to 'standard' philosophical doctrine. All good philosophers are 'unorthodox', i.e. critical of some view which they like to regard as 'standard', and the 'Double Aspect' theory of Body and Mind defended in this book is not entirely remote from such 'standard' theories as those of Spinoza and G. F. Stout. What a philosopher can contribute is the introduction of students of psychology to the awareness that some issues are essentially controversial. So many students come to college expecting to be told the Truth, and to be given only 'the established facts'. One established fact is that on so many questions the truth cannot yet be told. This can be demonstrated in the teaching of any subject, but it can be conveyed most expeditiously in the discussion of philosophical issues.

This book should be a prescribed text for all first-year students of psychology. It is also a challenge to all teachers in departments of psychology elsewhere. This is what, apparently, is taught to psychology students at Glasgow—what do you teach in your department?

Students who have covered the ground of *Human Senses and Perception* will be ready to tackle *Behaviourism and Phenomenology*—a model of interdisciplinary collaboration through the medium of the conference.

The volume belongs to a series of reports of symposia organized in 1962–63 to commemorate the fiftieth anniversary of the William Marsh Rice University. It is edited by Prof. T. W. Wann, Associate Professor of Psychology at Rice University. The other distinguished contributors are Sigmund Koch, R. B. MacLeod, Carl Rogers, B. F. Skinner and two of the most forceful representatives of contemporary philosophy in the U.S.A., Norman Malcolm and Michael Scriven.

Koch, under the title 'Psychology and Emerging Conceptions of Psychology as Unitary', reviews an important phase in the near-contemporary history of psychology. It is a panoramic survey of the kind which can only be enjoyed and assimilated into the 'cognitive map' of the reader. It is supported by a hand-picked bibliography of eighty-one items. MacLeod follows with a learned historical piece covering developments from the phenomenal*isms* of the British Empiricists to the phenomeno*logies* of Husserl and the Existentialists. The conjunction tends certainly to complicate if not to confuse the central issue. Carl Rogers in 'Towards a Science of the Person' develops the phenomenological and existentialist idea in a characteristic way.

It will not be surprising if many readers select Skinner's 'Behaviourism at Fifty' as outstanding among a collection of outstanding papers. After all, Skinner is the reigning Pope of Behaviourism, supported though he be only by a college of unreliable cardinals. Skinner understands better than most behaviourists the things that puzzle philosophers. He can speak the philosophers' 'language'. It is important that he agrees with philosophers such as Malcolm that Behaviourism is a philosophical, not a psychological theory.

Norman Malcolm, who has done so much to put across to psychologists the teachings of the later Wittgenstein, here tries to put across a linguistic type of solution of the issue—that it is a matter of the analysis of statements 'in the first person present tense'. Clearly, surely, it is a matter of the analysis of a certain subclass of such statements. Further discussion between Skinner and Malcolm may lead to a definition of the subclass in question.

Michael Scriven gives the concluding paper under the innocent title 'Views of Human Nature'. It is characteristic in its penetration, good sense and good humour.

An attempt is made in this book to record the discussions which followed each paper. What stands out between the gaps of inaudibility in these recorded discussions is the immeasurable enhancement of the standards of sophistication and courtesy in discussion today since the days of the slapstick argumentation of Watson and MacDougall. See, for example, the answer given by Scriven in the last session to 'A Question from the Audience': 'On behalf of students of psychology I should like to ask this: What is psychology?' See, indeed, the whole book. This volume, like *Human Senses and Perception*, is a quite essential book—at least for all departmental libraries of philosophy and psychology.

C. A. MACE

The Act of Creation. By ARTHUR KOESTLER. London: Hutchinson. 1964. Pp. 751. 42*s.*

This is a book that will, undoubtedly, repel many professional psychologists. The time is long since past, they will say, when a mere *littérateur*, even one as brilliant and versatile as Arthur Koestler, can make any useful contribution to their science. And, no doubt, with respect to large areas of psychology where exact experimentation has happily superseded armchair speculations such an attitude would be fully justified. But one has only to glance at some of the topics which the author covers—laughter and crying, creativity in the sciences and the arts—to realize how far this is from being the case in the present instance. Indeed, as anyone who has ever had to lecture to undergraduates on any of these topics knows only too well, to restrict oneself to the findings of experimental research means, in effect, to talk only about trivialities and incidentals. And whatever else one may say about this book it is not, like the great majority of books that claim to introduce a scientific approach to this field, full of pedantic irrelevancies. Considered just as a work of exposition, it has a verve and flourish that puts it into a class of its own and I should consider it a dereliction of duty if I did not encourage my students to read it.

The author, however, has made it very plain that he wishes it to be judged as an original contribution to psychology, and it is as such that we must now discuss it. Its thesis, which is essentially that which he put forward in his earlier and much more sketchy work *Insight and*

Outlook, is in very bald outline as follows. Each of our activities, whether those of the intellect or simple motor habits, is governed by a set of implicit rules or 'code', as the author calls it, and each code determines a particular 'matrix' which represents the range of possible operations that the code permits. Behaviour can be regarded as made up of a hierarchy of such relatively autonomous *matrices*. The development of this hierarchy involves two opposite processes: a downward process whereby acquired habits and skills form matrices that function unconsciously and with an increasing degree of autonomy, and an upward process whereby existing matrices combine to form new matrices and new codes. It is the latter process that constitutues the 'act of creation'.

Crucial to Koestlerian psychology is the concept of 'bisociation'. Two matrices are said to be *bisociated* when they have been brought together through some element or feature they have in common. When this happens a number of important consequences may ensue. Where the two matrices carry incompatible emotional charges the collision may issue in a burst of laughter. (Koestler's discussion of humour which occupies the whole of the first part is, I consider, the best contrived and most convincing part of the book). When, however, the 'participatory emotions' prevail over the 'self-assertive' ones, the effect may be tragic or aesthetic as the case may be, as when a great artist can make us see familiar things in a new light. But it is in the domain of science and discovery that this concept has its most far-reaching implications for it is then that we most need to be jolted out of the rut of habitual trains of thought that are leading nowhere. In this context bisociation becomes equivalent to that sudden act of insight about which psychologists have had so much to say. Here Koestler can bring to bear his great interest in the history of science, and among the passages I most enjoyed in this book (some of them relegated to appendices) were those where he discusses episodes from the lives of men like Darwin or Pasteur.

It would indeed be incredible if a work of this scope did not expose its author to some searching criticisms. Here I shall mention briefly what I regard as some of its more serious weaknesses. In the first place, the author has succumbed to the fatal temptation to overreach himself. When a theory purports to explain everything from molecular biology or the process of sexual reproduction to the higher flights of scientific and artistic imagination with the same few basic concepts, the reader may well begin to suspect that perhaps it really does not explain anything at all but is aimed merely at trying to dazzle him with clever but superficial analogies. Certainly on the Popperian criterion of falsifiability this would not qualify as a scientific theory since Koestler nowhere proposes experiments which, if they did not conform to his predictions, would make him retract or at least modify his theory. Nevertheless, as a plea for a new orientation in psychology much of what Koestler has to tell us makes excellent sense. Indeed, the concept of behaviour as a hierarchy of 'programmed' routines rather than a chain of conditioned reflexes is one that has already found much support since cybernetic models began to replace the older stimulus-response analyses.

A more serious fault is Koestler's curious failure to realize that in the study of thinking the initiative today rests with those who are concerned with the problem of artificial intelligences. To design a machine that can solve complex intellectual problems demands, after all, an understanding of the logic of problem-solving that far surpasses any merely verbal formulations, even if, in practice, the computer achieves its ends by other means than those exemplified in a human brain. The fact that Koestler can slur over this momentous development with a casual reference to machines that can play noughts-and-crosses shows how out of step he is here. The trouble, however, is not one of ignorance (Koestler's erudition is truly impressive, despite the occasional unexpected omission—there is, for example, no mention of J. S. Bruner), but the fact that he has let himself get distracted by his somewhat tangential polemic against the early behaviourists: Thorndike, Pavlov, Watson, etc. We are only too ready to grant him that human beings are not penny-in-the-slot machines; what we would like him to tell us is precisely wherein we differ from those sophisticated automata by means of which cyberneticists are now trying hard to simulate human behaviour and the higher mental processes.

Koestler harps continually on what he regards as the neglected unconscious aspects of thought, but where his argument becomes confusing is that he never makes clear what he really means by the unconscious. We are given to understand that we cannot equate it with the purely automatic working of the brain in the absence of conscious experience, but we are granted no positive conception of what it might be. A similar cloud of ambiguity is left surrounding the act of creation itself. In his zeal to destroy the mechanistic view of mind, Koestler will have no truck with any explanation that makes insight merely the fortunate end-product of some internal process in

which potential solutions are constantly being shuffled around. Yet, if we deny that thinking involves trial-and-error or competition-and-selection in some form and at some point, we are committed surely to a quite mysterious notion of the problem-solver being inexplicably drawn towards the correct solution by some sort of unerring intuition. Perhaps, when one is dealing with something as extraordinary as the human mind, mysterious and even occult notions may have their place (or, at least, I am hardly the one to cavil at them), but I do prefer an author who lets you know just where he stands.

JOHN BELOFF

Psychology: A Study of a Science. Vol. 6. Edited by SIGMUND KOCH. London: McGraw-Hill. 1963. Pp. xii + 791. 97*s*.

The sixth volume of this monumental work, subtitled 'Investigations of Man as Socius', is concerned with the relationship between psychology and other social-science disciplines. As with other volumes, the objective is not to provide information about the results of research, though to be sure, a great deal of information is incidentally included as illustrative material in every paper. The purpose is rather to discuss in general terms the kind of relationship which exists, or might exist. The question is: in what kinds of ways do other disciplines make use of psychological theory and findings?

There are fourteen papers, and all one can do is to give some indication of the topics covered. Muzafer Sherif starts the ball rolling with a general paper on 'Social Psychology: Problems and Trends', in which he points out how the social psychologist is concerned with the wider society, as studied by sociologists and anthropologists, and how the study, for example, of small groups can make a contribution to wider issues. Donald T. Campbell is concerned with 'Acquired Behavioural Dispositions', and especially with attitude formation. In his paper he points out the way in which the same phenomena can often be explained in terms of different psychological theories—for example perceptual selection on the basis of past experience or straight learning theory.

There are two papers on psycholinguistics, one by Charles E. Osgood, the other by F. G. Lounsbury. The subject-matter of linguistics is outlined, and attention is paid to the psychology of language learning. Both papers are extremely informative.

The relationship between psychology and anthropology is an obvious topic, and it is tackled by David French, who concentrates on the relevance to anthropology of psychological theories of perception and cognition, while George and Louise Spindler are concerned with culture change. The latter paper, in addition to some interesting material, includes an analysis of the literature on cultural change between 1929 and 1962 in order to find out to what extent anthropologists studying culture change have used psychological material. This short content analysis is unobjectionable at the end of a paper, but it is an unrewarding method when it is used to supply the material for a whole article, as is the case with William W. Lambert's discussion on 'Social Psychology in Relation to General Psychology and other Behavioural Sciences'.

Alex Inkeles contributes a paper on 'Sociology and Psychology', which ought to be recommended to any student of the former subject.

The way in which students of political science make use of psychological concepts and research is dealt with by Robert E. Lane. The relevant concepts are obviously leadership, public opinion, authoritarianism and methods of personality assessment, both of individuals and of collections of individuals.

There are four papers on economics and psychology. The first and, for the psychologist, the most rewarding is by George Katona; this is followed by a short comment by James Toban and F. Trenery Dolbear Jr.; a general paper, perhaps more interesting to economists, by Herbert A. Simon, and a sophisticated analysis of methods of assessing utility and expectation in economic behaviour. The point is made in the first and third of these papers that something that can be called 'economic psychology' has long been part of the equipment of economists, particularly when they are concerned with the verification or, more often, the falsification of the classical model and common-sense beliefs. Katona makes the point that economic behaviour ought to be studied more by psychologists; economics is a field in which hypotheses concerning motivation and attitude can be tested in measurable terms.

Finally, though it is not the final contribution, there is a curious paper by A. Irving Hallowell on behavioural evolution. The theme is: is there a gap between *homo sapiens* and all other animals?

This is widely assumed, but Hallowell suggests that a study of the behaviour of other animals, particularly primates, will reveal that there are adumbrations of behaviour that is thought of as specifically human, so that theoretically some kind of evolutionary continuum can be assumed in the behavioural field.

W. J. H. SPROTT

Readings in Psychological Tests and Measurements. By W. L. BARNETTE, Jr. Homewood, Ill.: The Dorsey Press. 1964. Pp. ix + 354. 24s.

This book of readings, intended to supplement basic texts, has several unique features which make it more than a collection of articles. Considerable effort has been expended in editing the original papers so that there is a consistency of statistical presentation and nomenclature, sometimes involving abbreviation or simplification. The aim of the extensive editing was to enable the volume to be useful to undergraduate students majoring in psychology. The scope and content ensure that the book will be of value to anyone taking a course in psychological testing at any level and to others working in the testing field who wish to refresh their ideas on key concepts in this field.

The forty-eight papers are grouped in eleven sections: general measurement problems, test administration problems, norms, response set, reliability, factor analysis, validity, intelligence, personality, interests and critiques of testing. In all, seventeen papers relate to validity while all other sections have between two and five each. With few exceptions, all papers are post-1950, thirty having been published in 1956 or later. Papers by Thurstone and French on factor analysis and by Flanagan on the experimental evaluation of a selection procedure are among the earlier ones. For each paper there is an editorial introduction in which the background to the problem is set out and, where appropriate, an indication of what knowledge is needed before the paper can be appreciated. For example, it is noted that before reading a report by Worbois on the prediction of long-range performance, the reader should know the basic facts of multiple correlation and statistical weighting technique and of multiple critical score technique—a statement which is immediately followed by the name of a non-technical source of the information required. In other cases, the introduction also has a brief discussion of other papers related to the topic, some of which have been referred to by the original author in his paper.

Readers already familiar with the field of test application may feel that there are other papers which might have been included. However, most readers will agree that the papers in this volume can but aid the beginner towards a clearer appreciation of the many problems that can beset the worker in the field of psychological testing and measurement.

K. M. MILLER

Auto-primer in Computer Programming. By DORIS R. ENTWISTLE. New York: Blaisdell. 1963. Pp. 345. $6.50.

This auto-primer is an excellent introduction, in the form of a linear programmed text, to programming in FORTRAN, one of the several computer languages currently in use. To those who have access to an IBM computer or, in fact, to any computer with a FORTRAN compiler, this text is thoroughly recommended, whether as a basic introduction to the topic of programming, as a refresher for those whose programming techniques may have degenerated with misuse, or as a supplementary guide for use in conjunction with a formal programming course.

This text should also prove a useful introduction to programming techniques for those who currently have no specific compiler in mind but who would like to learn something about computers and programming. The possession of some degree of skill in FORTRAN programming will be unlikely to be wasted, as this is probably the most widely used computer language.

This book should serve as a model for introductory texts on the subject of programming. Aimed basically at graduates, and assuming only the most elementary mathematical knowledge, it moves at just the right pace, is clearly presented, and contains enough supplementary information to allow the reader to gain more than a simple mechanical competence in programming.

R. S. CORTEEN

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION. *The Scottish Scholastic Survey, 1953.* London: University of London Press, 1963. Pp. 216.

There is singularly little of psychological interest in this, the latest of the Scottish Research Council's excellent publications. It describes the construction, application and results of tests of Mechanical and Reasoning Arithmetic, English Usage and Reading Comprehension, given to almost the complete Scottish primary school population aged 9 years 11 months to 10 years 10 months (totalling over 72,000) in May 1953. The tests were similar to the Moray House 11+ series, and clearly involved careful sampling of the primary school syllabus. But as they were not tried out properly beforehand, the score distributions are oddly skewed. However, they attained high reliability coefficients and, as detailed norms are provided, they should be useful for future testing within this age range.

Most of the book is taken up with analysing the results on various test items, and pointing out the implications for the teaching of English and arithmetic. The following are the main conclusions of more general interest. When children of this age do rough work with their sums, they are more likely to get them right. The usual sex differences emerge—boys better at Arithmetical Reasoning, little difference in Mechanical, girls better at both English tests; the boys show somewhat greater dispersions in all tests. There are some geographical differences between ten main areas of the country ranging, apparently, up to nearly $\frac{1}{2}\sigma$. Left-handed *writers* totalled 6·8% of boys and 5·1% of girls; their mean attainment was only slightly inferior to that of the right-handed. No regular relation could be found between attainment and size of school class, but on the whole 2 to 5 teacher schools did a little less well than 1 teacher of 6+ teacher schools (possibly this merely reflects regional differences).

The work is described with admirable clarity but one wonders whether, in the absence of any definite hypotheses to be investigated, it was worth all the trouble.

P. E. VERNON

Creativity and Psychological Health. By FRANK BARRON. New York and London: Van Nostrand. 1963. Pp. x+292. 51*s*.

Creativity: Progress and Potential. Edited by CALVIN W. TAYLOR. New York and London: McGraw Hill. 1964. Pp. xi+241. 54*s*.

Education and the Creative Potential. By E. PAUL TORRANCE. Minnesota and London: Minnesota Univ. Press and O.U.P. 1963. Pp. vii+167. 36*s*.

In recent years research into 'creativity' has expanded rapidly in the U.S.A. in spite of some scepticism as to whether it is a useful or a unitary concept. Whatever the answer to the latter question, this renewed interest has resulted in a number of diverse approaches. Creativity has been described in terms of unconscious and preconscious processes (Kubie), as a form of 'divergent' thinking (Guilford), as largely overlapping with general intelligence (Burt, Vernon) and as predictable from a combination of primary personality factors (Cattell).

The three writers under review assume that creativity is an important topic in its own right, but approach it from different angles. Barron's book will perhaps be of greatest interest to the general and clinical psychologist, Taylor's to the research worker in the field, Torrance's to the educational psychologist and the teacher.

Barron ranges widely and speculates freely, and this breadth of treatment is illustrated by some of the chapter headings, 'Signs from psychological tests that promise improvement in psychotherapy', 'Rebelliousness, morality and psychological health', 'Intelligence quotient, personality and originality', 'Unconscious and preconscious influences in the making of fiction'. Empirical work by the author and his colleagues is described, in particular an elaborate study of 'personal soundness' in graduate students, and another of originality and its correlates in U.S. Air Force officers. The book may sound a hotch-potch from this description. It is not, but its unity is a personal one, deriving, as stated in the preface, from the author's first introduction to psychology, his experiences during the war and his religious beliefs, as well as from

academic interests. The result is a valuable and humane contribution to psychology, much of which will be intelligible and interesting to the educated layman.

Whereas Barron was mainly concerned with the springs of original thinking in the individual and more with literary than with scientific creativity, Taylor's contributors summarize work on the prediction, identification, education and encouragement of original talent, particularly scientific talent, in the community. The aims, competently fulfilled, are to state plainly what is known in this whole area and to indicate urgent lines of research and action. These include, for example, the investigation of pre-verbal and non-verbal manifestations of creative thinking in very young children, and a re-evaluation and re-direction of teacher training to encourage creative thinking. This is a useful reference book and contains a valuable bibliography of some 700–800 items.

Torrance presents a number of addresses given to teachers, psychologists and others, supplemented by experimental studies, the main common theme being the encouragement of original and creative thinking in schoolchildren. The result is a rather scrappy book, but one containing much that is, or should be, of interest to educational psychologists. Torrance's findings are optimistic in one respect, pessimistic in another. He is convinced, for example, that primary school children can, without great difficulty, be taught to think more creatively, but throughout the book there is abundant evidence of the obstacles and negative pressures that frequently arise from the attitudes of parents and other children. The most depressing of these seem to be typical of American ardour for conformity, but similar research in this country would probably leave little room for complacency.

H. J. BUTCHER

Clinical Psychology: An Introduction to Research & Practice. By N. D. SUNDBERG and L. E. TYLER. London: Methuen. 1963. Pp. xvii + 564. 50*s.*

This is a sound book but rather a dull one, full of truisms whose omission would have left it shorter and a lot more readable. The present reviewer, while respecting Professor Tyler's previous work on individual differences, must admit to an initial bias against the American interpretation of clinical psychology. Respect for the authors' work, however, increased with reading.

The book would probably make a better introduction to counselling than to clinical psychology; and for British psychologists the three introductory and the three concluding chapters have little relevance, since they reflect so largely the present position of clinical psychology in American culture. In the introductory chapters the authors attempt to come to conceptual grips with their subject-matter. Their preference is for a developmental framework, which should be congenial enough to British students. The final chapters, which are concerned with the role of the clinical psychologist in society, are again largely irrelevant to British conditions.

The remaining three-quarters of the book, which deal with psychological methods of assessment and treatment, contain much that may be useful for the postgraduate student of clinical psychology, and even for the undergraduate considering clinical psychology as a career. Sound advice is given on such practical matters as how the psychologist relates himself and his work to the patient. At the same time, the authors summarize a large amount of research related to clinical psychological problems, giving due emphasis to the studies of Meehl and Sarbin, and to others inspired by them, on the rather poor validity of clinical prediction. This smooth welding into the body of the text of so much relevant research is probably the authors' greatest accomplishment. They bring out also the *ad hoc* 'do it yourself' nature of much of the clinical psychologist's day-to-day work. The problem of his ambivalence, of how he is caught between his clinical impressions and his test results, is fairly stated. There is some honest discussion of how much of present-day clinical activity, especially on the assessment side, is really useful; the authors seem to agree with Meehl's earlier conclusion that much of it may not be. Short reviews of different psychological methods of treating mainly the psychoneuroses give the reader a quick and objective acquaintance with what is available in the field of psychotherapy, and would be particularly helpful to new graduates about to embark on a course of clinical training.

The list of 'Fifty tests of importance in clinical psychology', contained in Appendix A, would be neither adequate nor up to date so far as work in this country is concerned. The bibliography, on the other hand, is extensive and represents a very ample coverage of the American literature up to about 1961.

J. GRAHAM WHITE

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology

The British Journal of Psychology (Monograph Supplements)

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology (about 512 crown quarto pages per volume), *The British Journal of Medical Psychology* (about 350 crown quarto pages per volume), and the *British Journal of Social and Clinical Psychology* (about 320 crown quarto pages per volume) appear in quarterly parts, the subscription price of each, payable in advance, being £4. 10*s*. per volume post-free ($16.50 in the U.S.A.). Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, N.W. 1. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 288 pages. The subscription rate, payable in advance, is £3. 10*s*. per volume post-free ($13 in the U.S.A.). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, London, E.C. 4; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number. The subscription price in the U.S.A. is $4.20. Subscriptions should be sent to the Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36 Essex Street, Strand, London, W.C. 2.

A series of *Monograph Supplements* is issued in connexion with *The British Journal of Psychology* but is not included in the subscription for that journal. Orders may be sent to any bookseller, or the Cambridge University Press, 200 Euston Road, London, N.W. 1, from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor A. Summerfield, Department of Psychology, Birkbeck College, Malet Street, London, W.C. 1;

for *The British Journal of Medical Psychology* should be sent to

Dr T. Freeman, Lansdowne Clinic, 4 Royal Crescent, Kelvingrove, Glasgow, C. 3;

for *The British Journal of Social and Clinical Psychology* should be sent to

Mr Michael Argyle (Social Psychology Editor), Institute of Experimental Psychology, 1 South Parks Road, Oxford, *or* to

Dr R. R. Hetherington (Clinical Psychology Editor), 6 Abercromby Square, Liverpool 7;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Dr R. J. Audley, Department of Psychology, University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Mr L. B. Birch, The Institute of Education, The University, Sheffield 10.

for *The Monograph Supplements* should be sent to

Professor I. M. L. Hunter, Department of Psychology, The University, Keele, Staffs., whose decision regarding the publication of all Monograph Supplements will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1 September of the preceding year.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £6. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1 annually), *The British Journal of Medical Psychology* (£1 annually), *The British Journal of Social and Clinical Psychology* (£1 annually), *The British Journal of Mathematical and Statistical Psychology* (£1. 10*s*. annually), and *The British Journal of Educational Psychology* (£1 annually).

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, Tavistock House South, Tavistock Square, London, W.C. 1.

Preparation of Papers for the British Journal of Psychology

In submitting papers for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references, and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. 56 (1965) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.

Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* 26, 231–48.

GALTON, F. (1883). *Inquries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *L'Année psychol.* 50, 83–96.

World list abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

For other general guidance, authors are referred to *The Journal of Physiology: Suggestions to Authors*. Cambridge University Press. Price 2*s*.

OCCUPATIONAL PSYCHOLOGY

Editor: ALEC RODGER

JANUARY 1965

Volume 39, No. 1

CONTENTS

APRIL 1965

Volume 39, No. 2

CONTENTS

ARCHIVIO DI PSICOLOGIA NEUROLOGIA E PSICHIATRIA

YEAR XXV. ISSUE 5 SEPTEMBER–OCTOBER 1964

CONTENTS

LEADING ARTICLE

YEAR XXV. ISSUE 6 NOVEMBER–DECEMBER 1964

CONTENTS

LEADING ARTICLE

YEAR XXVI. ISSUE 1 JANUARY–FEBRUARY 1965

CONTENTS

LEADING ARTICLE

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor: I. K. WATERHOUSE, *University of Melbourne*

VOL. 17, NO. 1 APRIL 1965

ARTICLES

VOL. 17, NO. 2 AUGUST 1965

The Journal is published three times a year. Subscription rates 36*s.* (sterling), and $5.50 in U.S.A. Members of B.P.S. and A.P.A. should enquire about special subscription rates from the Secretary of their society. Non-members should send enquiries to Melbourne University Press, Parkville, N. 2, Victoria, Australia, or to its agents

in U.K.
Cambridge University Press
200 Euston Road,
LONDON, N.W. 1.

in U.S.A.
Cambridge University Press
32 East 57th Street,
NEW YORK, N.Y. 10022

Enquiries from the Orient should be sent directly to Melbourne University Press.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Edited by O. L. ZANGWILL

Vol. XVII, Part 1 February 1965 20*s*. net

Vol. XVII, Part II May 1965

Published for the Experimental Psychology Society by

W. HEFFER & SONS LTD, CAMBRIDGE

Annual Subscription, 60s. to W. Heffer & Sons Ltd, or any bookseller

JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR

A NEW SERVICE FOR JEAB SUBSCRIBERS

As the research literature on behavioral processes continues to accumulate at an ever-increasing pace, the task of locating previous material on a given topic becomes correspondingly difficult, frustrating and time consuming. As a step toward meeting this problem, the *Journal of Experimental Analysis of Behavior* proudly announces a new service that has been enthusiastically received by its subscribers. In addition to its volume index, the Journal is now distributing standard 3 in. × 5 in. reference cards for all articles and notes published in its pages. These cards provide suggested headings for appropriate classification of each item, with space for the reader to add his own. The cards also include the author, title, volume, pages, and an abstract of the contents.

Subscribers to the Journal receive these cards free of charge.

Subscriptions:	**$16.00 regular**
(calendar year)	**$9.50 individuals**
	$8.00 students

Please address subscription orders to:

KAY DINSMOOR, JEAB
DEPARTMENT OF PSYCHOLOGY, INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405

A price list of back issues is available on request

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 121–133
Printed in Great Britain

SOME STUDIES OF PATTERN PERCEPTION USING A STABILIZED RETINAL IMAGE

By C. R. EVANS*

J. J. Thomson Physical Laboratory, University of Reading

Using a stabilization device capable of covering 30° of the visual field, targets of various shapes were studied extensively. Very large and consistent differences were found between the percentage disappearances of most targets and this was found not to be a function of length of line or boundary. The presence of corners and intersections in a target increased percentage disappearance. Jagged, angular figures disappeared more than rounded, topologically similar ones. When acute angles were present, the disappearance rate was radically raised.

About 10% of the time when a target disappeared, it disappeared as a complete unit, though this figure varied somewhat from target to target and was very high in the case of the circle. A mean of about 15% of all disappearances were 'patterned' or 'structured', and these proportions held true for all subjects tested.

It is now well established that an image stabilized on the retina disappears intermittently. Previous workers have observed in addition that certain images seem to disappear more readily than others, and that any one pattern does not necessarily disappear as a whole (Ditchburn & Pritchard, 1956; Pritchard, Heron & Hebb, 1960). These observations have however been entirely qualitative and subjective. In the investigations reported here, systematic quantitative measurements have been made for the first time, and the results suggest that disappearances of these images cannot be regarded as a purely retinal phenomenon. Some central mechanism seems certain to be involved.

The work falls into two parts. In the first the hypothesis that targets of one shape disappear more frequently or more readily than another has been examined. In the second the apparently non-random character of the fragmentary disappearances of any one target has been considered.

Method

(*a*) *The stabilizing system.* The simplest method of producing a stabilized image involves attaching the target to a contact lens worn by the subject. The system employed in the present investigation is illustrated in Fig. 1. A full-size haptic contact lens moulded to fit the subject's eye is worn. Attached to the centre of the lens over the corneal bulge is a light-weight aluminium tube, into which fits another, of fractionally smaller diameter, which contains the focusing lens; a 'gate' or 'trap' on the end holds the target, which can thus be brought into sharp focus. The gate is constructed to allow small adjustments to be made in the position of the target. The weight of the entire system (less the contact lens) is slightly under 0·8 g. This system offers certain positive advantages over those employed in previous investigations, the most notable perhaps being that illumination of the image is constant and uninterrupted throughout the experiment, and eye blinks, etc., do not occlude the image.

(*b*) *Method of illumination.* Each subject lay on his back on a couch with head and shoulders encased in a relatively light-proof and sound-proof cubicle. At a point about 2 ft. above him was a large screen, evenly illuminated from behind by a battery of 'cold light' bulbs. The light passed through two homogeneous screens before reaching the subject and produced a diffuse light area of about 3 ft. × 2 ft., against which the target was viewed. No restriction upon the subject's head or body movements was necessary, for all normal movements produced no apparent

* Now at Autonomics Division, National Physical Laboratory, Teddington.

change in field brightness or stimulus contrast. The subject's other eye was occluded. Field luminance could be adjusted by the experimenter, through the use of a variable voltage transformer, from zero to 2 log. foot-lamberts.

(c) *Stimulus material.* Patterns drawn in black ink were photographed on 35 mm. positive film; i.e. they appeared as black targets upon a white ground. The average target size in the reduced state was about 3 mm., which, when viewed in the apparatus, subtended a visual angle of about 5° of arc in a monocular visual field extending over 32°. With the exception of the target itself, the perceived field appeared to be of even brightness.

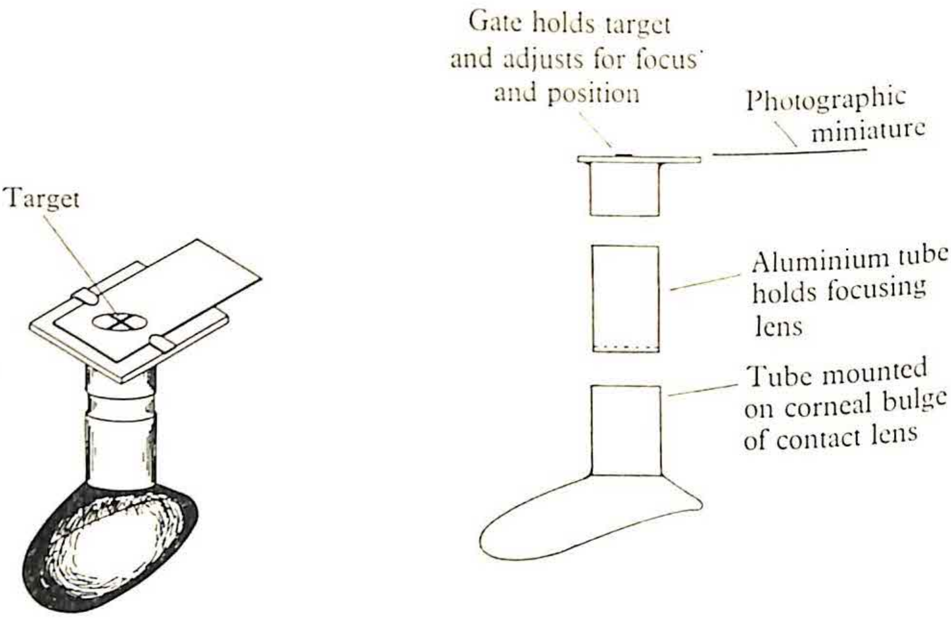


Fig. 1. The stabilizing system: the subject lay on his back, viewing the target against a diffusely lit screen.

Before the start of an experimental run the target was adjusted by the experimenter until the subject reported that it was in clear view and foveally sited, i.e. that he was fixating the centre of the target. There was no difficulty in making such adjustments, even with the apparatus in the subject's eye, but the gate nevertheless held each target so securely that its position, once set, did not alter in the course of an experiment.

Observers of stabilized images in apparatus of this kind noted that a target which was out of focus was a particularly irritating experience, and not one which could be tolerated for long. Constant attempts to accommodate were made and, when these proved fruitless, subjects reported restlessness and annoyance. Equally irritating in stabilized vision were targets which were off centre; they provoked quite involuntary and highly persistent tracking movements in an effort to centre the image. The subjects themselves therefore tended to insist on correct centring and focusing before starting an experimental run.

(d) *Method of Recording.* Two electrically driven counters registering at the rate of 250/min. were used, one being set to operate continuously throughout a trial, the other operating only when, and as long as, the subject pressed a Morse key. Thus, at the end of a minute's trial one counter would register 250, and the other a figure representing the amount of time for which the subject had pressed the key. In most cases subjects were instructed to press the key for as long as some part of the image had disappeared, and this could quickly be expressed as a percentage of total viewing time.

(e) *Subjects.* In any experiment involving attachments to the eye subjects are obviously not easy to find. Three subjects participated in the studies reported here, the bulk of the observations being recorded with one subject, a computer operator on the staff of the Laboratory who was unaware of the nature or implications of the experiment. A very large number of trials were undertaken with each subject and a high measure of agreement between individuals was found.

(*f*) *Procedure.* Most experimental runs consisted of ten 1 min. trials, preceded by one test trial. Subjects were found to be able to tolerate such a period without showing signs of restlessness or boredom. A complete run, including inter-trial breaks, never lasted more than 20 min.

Experienced subjects, within seconds of settling down in the apparatus, will notice changes in the stimulus material. These changes vary from time to time, but the general trend is for there to be an almost immediate reduction of contrast between the target and its background. This is followed, again within seconds, by a distinct, and at first surprising, disappearance. This disappearance is occasionally of the complete figure, but is most often of some part or parts of it. Unless some radical change takes place in the subject's state, such as the onset of a fit of coughing or blinking, the disappearance effects continue to occur steadily throughout the course of a session. Reappearances either of the image as a whole, or of *parts* of the image, occur at intervals of from 1 to 5 sec.

Throughout this series of experiments subjects were instructed to report 'disappearance' by pressing the key only on those occasions when some part or parts or the whole had definitely vanished. They were told to disregard occasions when the image merely became hazy, lost contrast with its ground, etc.

The coming and going of parts of a figure is striking, but sometimes puzzling and confusing. Even experienced subjects who viewed the range of phenomena with dispassionate interest occasionally must have made errors of over or under-recording. These are assumed to have been randomly distributed over the large number of trials, and to affect the results only as background 'noise'. Because of this noise, and because subjects exhibited some day-to-day variations in results, trials were often spaced out over days or, occasionally, weeks. For any given target at any one brightness level eighty 1 min. trials were employed, requiring from ten to twenty experimental sessions.

Results

Comparisons of different shapes

Simple figures were chosen for the preliminary studies, a large sample being shown in Fig. 2*a–k*. Subjects viewed patterns for sessions of eleven 1 min. trials (the first in a session being a 'test' and not included in any of the following results). The following instructions were given: 'Press the key and keep it pressed as long as some part, parts, or the whole of the image have disappeared.'

Results from the first subject, *A*, for the set of figures based upon the circle (Fig. 2*a–d*), are shown in Fig. 3. Each individual symbol on the graph represents the mean of five 1 min. trials. With this subject the addition of a single bar to the circle raised mean disappearance considerably, and the differences between the means are highly significant. Similarly, the addition of a further line, which converted the target to a circle with an inscribed cross, caused a further jump in percentage disappearance. Again, the differences between means are highly significant. With Fig. 2*c* rotated through 45° to form Fig. 2*d*, in which diagonals were presented to the observer instead of vertical–horizontal lines, no significant difference in disappearance totals is found. Fig. 4*a*, *b* shows additional studies of the first three targets with two other subjects, B and C. Once again differences between overall means are highly significant.

At this point it was already evident that there is a major difference between the perceptual stability exhibited by figures of different shape. It seemed possible that this was a simple function of the length of line actually projected on to the retina. The circle with one line represents an increase in total length of about one-third over the circle alone, and the circle with cross is over 50 % greater than the circle. One might argue 'the more line, the more there is to disappear', but this simple hypothesis fails when the results from the next set of figures are examined. Here patterns without the circular component were tested, Fig. 2*e–g* being chosen. Width of line,

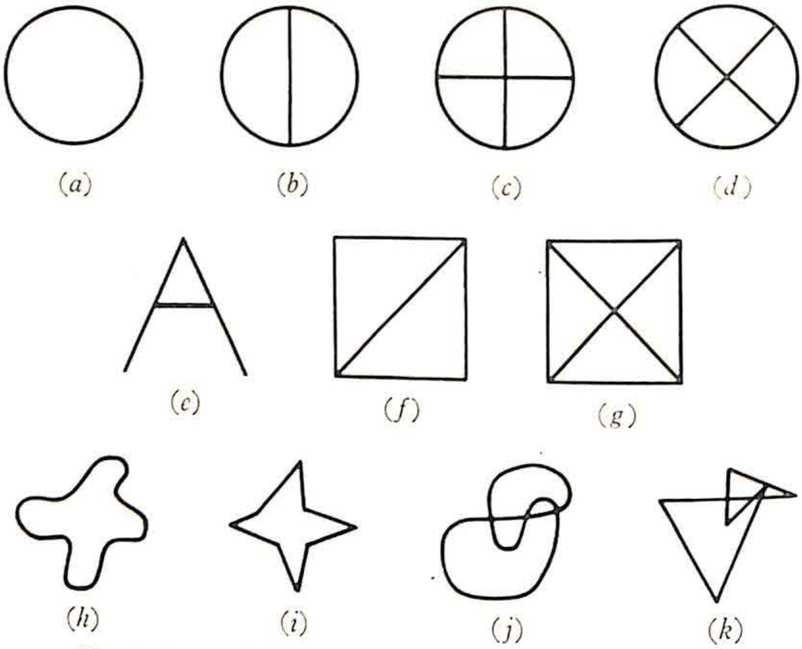


Fig. 2. Some of the targets used to give stabilized images.

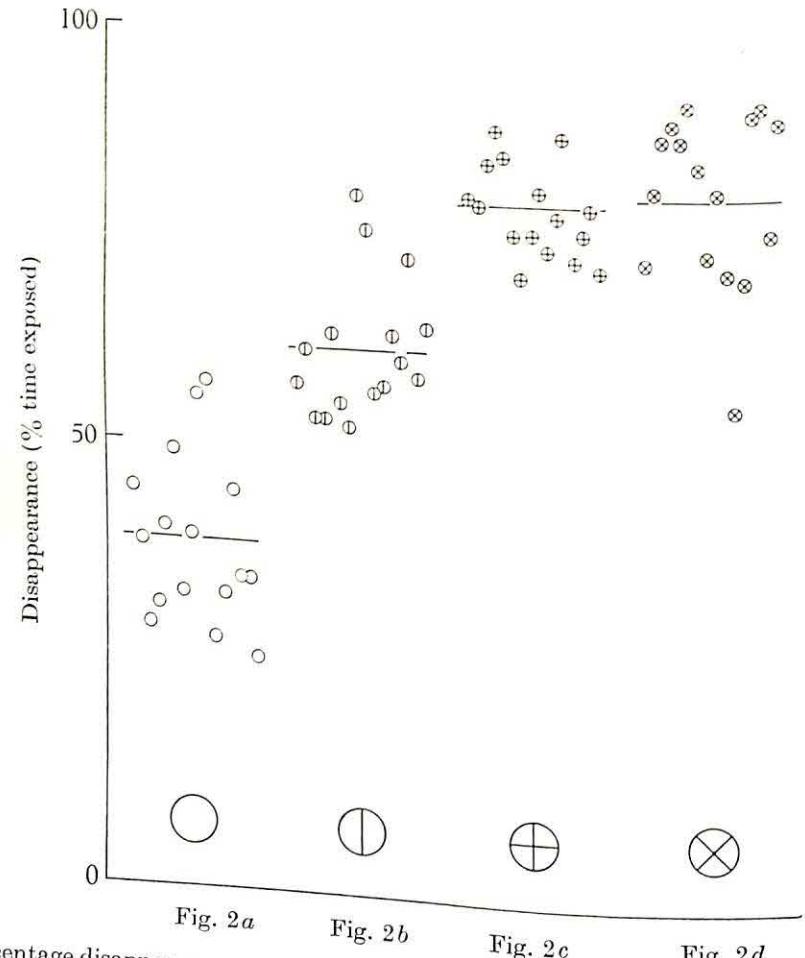
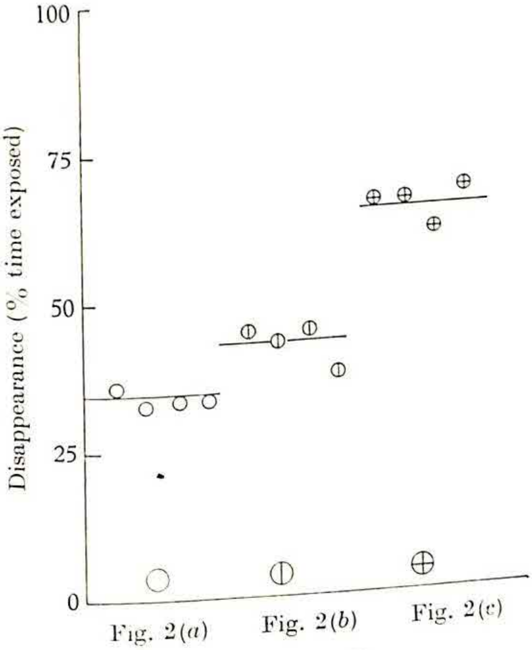


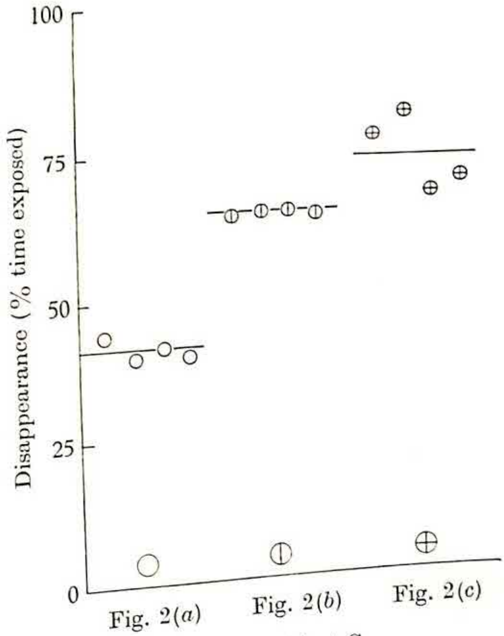
Fig. 3. Percentage disappearances of four different targets for subject A. Each symbol represents the mean of five 1 min. trials and the horizontal lines the overall means for each of the four targets, Fig. 2*a* (40·0), Fig. 2*b* (63·0), Fig. 2*c* (80·0) and Fig. 2*d* (81·0).

field brightness, contrast, and retinal projection area were kept constant with respect to the previous set.

Fig. 5 shows the results obtained with subject A for these patterns. Once again each symbol represents the mean of five trials. Differences between the means are highly significant and it is clear that disappearance values were not simple functions of the length of line projected upon the retina. Fig. 2*e* for instance has a consistently high percentage of disappearance, 76 %, which is nearly *twice* that of the circle, yet its retinal projection in terms of line length is actually less.



(*a*) Subject B.



(*b*) Subject C.

Fig. 4. Results for subjects B and C for comparison with those given in Fig. 3 for subject A.

At this point change of shape may mean change of disappearance rate or perceptual stability, but it is not yet clear what the effective variables are. There is some negative evidence however: length of line is relatively, if not totally, unimportant; moreover, it is not diagonals *per se* that cause the high rate of disappearance (cf. the lack of any difference between the patterns of Fig. 2*c*, *d*). One variable that occurs in the non-circular targets is the presence of acute angles. Fig. 2*e* has one at least, Fig. 2*f* has four, and Fig. 2*g* has at least eight. Pritchard *et al.* (1960) reported the impression that 'spiky' figures seemed less stable than 'rounded' ones, and they related this to the Gestalt concept of 'good figure'. This observation might not be surprising if acute angles themselves in some way bring about high disappearance rates, and this hypothesis may be tested by selecting two figures of equivalent area, length of line etc., one containing acute angles, the other none. Two satisfactory figures are the capital letters Z and S; Fig. 6 displays the result of such a test, each symbol and value again representing the mean of five trials. The results on the left were obtained from subject A, and on the right from subject B. Field luminance was set at 1·1 log. foot-lamberts. This luminance was somewhat higher than for the previous figures and the

increase was introduced in an attempt to depress disappearance values somewhat; the proximity of the values for Fig. 2g to 100% seemed to recommend this course. It is clear that the results are in accord with the hypothesis that acute angles are a factor in provoking high disappearance values; the differences between means are highly significant in the predicted direction. It might be objected that one important difference between the last two figures has not been controlled, namely, that they differ along the continuum of familiarity, S being a much more common letter than Z; some further check is required. Since familiarity could be an important variable, it is desirable to employ nonsense figures. Fig. 2h is a nonsense figure of vaguely amoeboid shape; Fig. 2i is its spiky analogue. The prediction, on the acute angle hypothesis, is that the spiky figure will disappear more than the rounded one. Columns (1) and (2) of Table 1 show the five-trial mean values and the highly significant difference between overall means that occurs in the predicted direction.

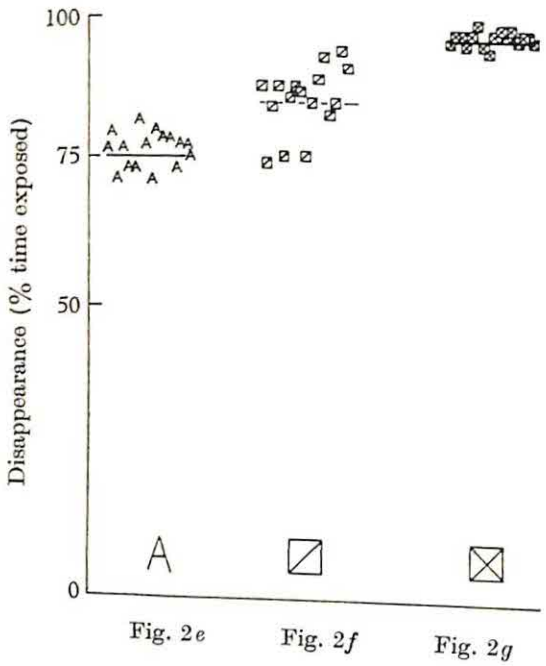


Fig. 5

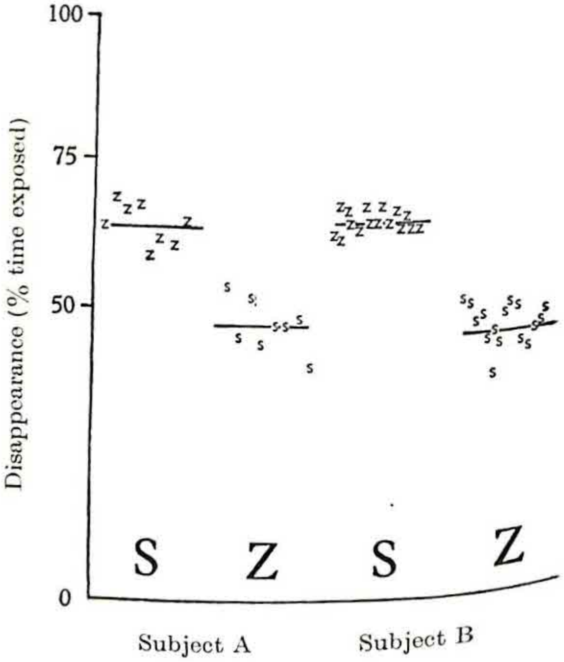


Fig. 6

Fig. 5. Percentage disappearances of three different targets for subject A. Each symbol represents the mean of five 1 min. trials and the horizontal lines the overall mean for each of the three targets, Fig. 2e (75·0), Fig. 2f (85·0) and Fig. 2g (96·0).

Fig. 6. Percentage disappearances of the capital letters Z and S used as targets for subjects A and B. Subject B made twice as many trials as subject A. Each symbol represents the mean of five 1 min. trials and the horizontal lines the overall mean for each target for each subject: for subject A, Z (62·0) and S (48·0): for subject B, Z (64·0) and S (46·0).

If intersections between lines (which may often form acute angles) are important, then a spiky nonsense figure with intersections (Fig. 2k) should disappear more than a rounded nonsense figure with intersections (2j). Columns (3) and (4) of Table 1 show this prediction to be borne out for subject A.

Before passing to the second part of the experiment, some attention might well be paid to the Gestalt point of view, since the results so far recorded would have been

viewed with much satisfaction by the pioneers of that school. While it may be true that it is acute angles that in some way trigger off high disappearance, it may equally well be true that 'good' figures, i.e. rounded ones, are inherently more stable than 'less good', spiky ones. If this seems to be saying nothing new, another Gestalt

Table 1. *Comparison of percentage disappearances of 'rounded' targets (Fig. 2h, j) and 'spiky' targets (Fig. 2i, k)*

(The differences between the means of columns (1) and (2) and of columns (3) and (4) are significant ($P < 0{\cdot}001$.)

	(1) Fig. 2*h*	(2) Fig. 2*i*	(3) Fig. 2*j*	(4) Fig. 2*k*
	58·0	74·9	66·2	88·9
	51·7	68·6	73·6	84·2
	55·5	75·0	82·8	75·1
	66·0	74·9	68·1	77·6
	56·4	81·2	76·2	84·1
	61·5	84·2	71·9	89·2
	58·6	82·2	61·0	79·1
	56·3	79·5	77·7	82·1
	48·1	78·5	83·7	83·1
	51·6	74·3	78·9	74·8
	40·5	72·7	64·6	86·7
	49·9	79·4	57·7	86·7
	56·8	84·1	66·1	96·6
	56·7	77·7	59·6	81·8
	65·7	80·8	70·2	81·1
	76·6	79·9	51·9	86·4
Mean	56·9	77·4	69·4	83·6
S.D.	12·1	7·2	12·0	8·6

Table 2. *Comparison of percentage disappearances for a circle and an ellipse of equal area and angular dimension*

(The difference between the means is significant ($P < 0{\cdot}001$), showing that the circle was less prone to vanish.)

	Circle	Ellipse
	46·3	52·4
	44·6	46·7
	38·3	52·1
	34·0	54·9
	48·1	58·0
	46·4	60·3
	42·3	57·5
	46·1	64·3
Mean	43·3	55·6
S.D.	9·9	9·9

prediction, in which angularity does not arise, may be proposed. All Gestalt theory suggests that an ellipse is a less good figure than a circle. If 'goodness' brings about resistance to disappearance in the stabilized image, one would expect an ellipse to disappear more than a comparable circle, and Table 2 shows the Gestalt prediction to be upheld.

The accuracy of subjects' recordings

While consistency within any given series and repeatability of results after a lapse of time are themselves good indications that a subject made careful and objective recordings, it is still of interest to have some independent check where possible. Since the phenomena of stabilized vision have no long history of study behind them, the choice of such an independent test is limited. Fortunately, from the collection of data gathered over the past decade there are a number of facts upon which all experimenters are agreed, and one of the most satisfactory of these is the finding that, as field brightness is raised, disappearances of an image become less frequent (Fender, 1956; Pritchard, 1958; Clowes, 1959), though this effect has never been studied for patterned targets. With this in mind, the principal subject, A, was tested in the following way. A triangle, of similar area, etc., to the targets previously described, was presented under the usual conditions of stabilization. The subject viewed the target, recording disappearances in the usual way, but at seven different levels of field luminance, with forty or more 1 min. trials at each level. Luminance levels were presented in random order.

Fig. 7 shows that disappearances fell off as a function of field luminance. The consistency of this graph suggests a high degree of accuracy and objectivity on the part of the subject, who was unaware of the nature of the experiment and was not informed that luminance was being manipulated as an experimental variable.

Partial and non-random disappearances

When stabilization was good and when targets were clear and sharply contrasted with their background, disappearance effects were almost immediate and very striking. In addition, all subjects who viewed a meaningful pattern under such conditions swiftly reported that disappearances were mostly partial, and that the fragmentation effects were not haphazard or random. What is meant by 'partial' and 'non-random' or 'structured' disappearances is illustrated by Fig. 8. This figure shows a target which was viewed many times by all three subjects as a stabilized image, and Fig. 8*a* a selection of possible 'structured' configurations which commonly occurred and often remained for periods of seconds. Fig. 8*b* gives four of a large number of possible configurations that may be called 'random'.

In some pilot experiments subjects seemed to have the impression that almost all disappearance configurations were structured, but this conviction appeared to dwindle with practice in observing. To counter any tendency to overestimate this category of disappearance, extremely rigid criteria were set for the scoring of a structured disappearance. Before the experiment subjects were shown a number of drawings of possible fragmentations of patterned targets. They were told to take great pains to report a structured configuration only when one of these clearly delineated figures occurred and persisted. If odd, random parts of the figure remained, even in association with some obviously non-random shape, they were not to report a structured configuration.

In an effort to determine the relative proportions of 'structured', 'random' and 'unitary' (i.e. disappearance of the figure as a complete unit) scores, a slight modification was made in the experimental procedure. In one experimental session subjects

were given six trials (including one test) for which the instructions were as before: 'Press the key and keep it pressed as long as some part, parts, or the whole of the image have disappeared'. After these six trials had been completed, a further five trials were given in which the instructions were: 'Press the key and keep it pressed only when the figure has disappeared as a complete unit and as long as it remains disappeared as a complete unit'. This, the first condition of the experiment, gave a record of the percentage disappearance of a given figure as a complete unit: it is emphasized that this is the percentage of time for which an object, *when disappearing*, does so in unitary fashion.

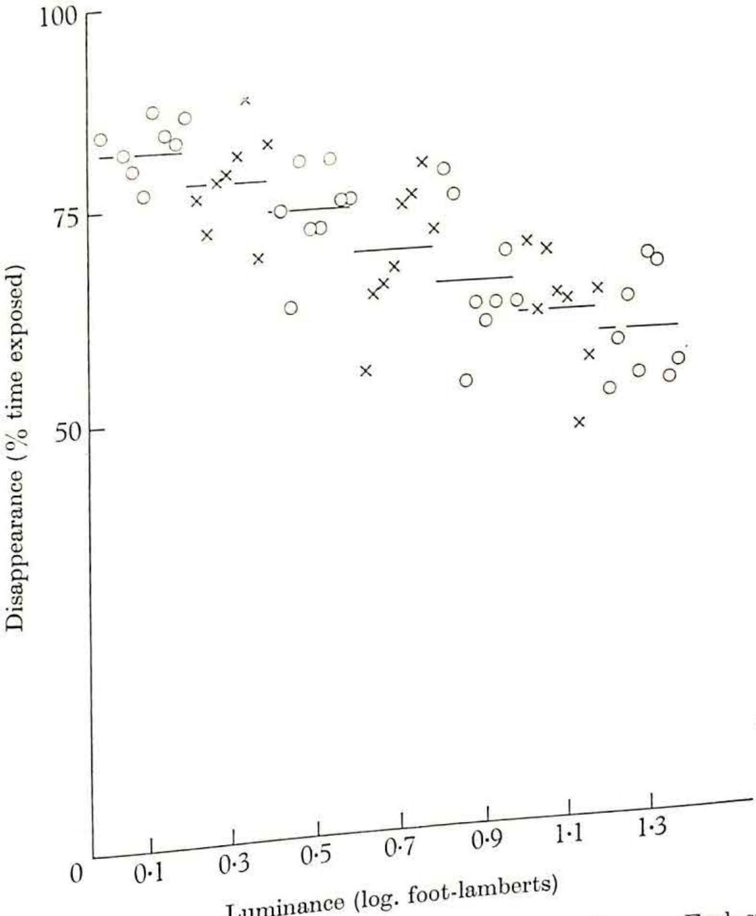


Fig. 7. Percentage disappearance falls off with increasing field luminance. Each symbol represents the mean of five 1 min. trials and horizontal lines the overall means for each of the seven values of field luminance.

In the second condition of the experiment (always constituting a new session), after the usual six 'all disappearance' trials, the following instructions were given: 'Press the key and keep it pressed when you get random, formless, or haphazard configurations persisting, and only as long as you are getting them.'

In the third condition, after the usual opening six trials, a further five were given with the instructions: 'Press the key and keep it pressed when you get structured configurations persisting, and only as long as you get them.'

With data collected under these conditions, a complete breakdown of disappearances into one or another of the above categories was possible, and a range of targets was studied. Fig. 9 shows percentage disappearances of each kind for subjects A, B and C when one target, the circle with an inscribed cross, was viewed for 160 1 min. trials per subject. It will be seen that the breakdowns show a very high measure of agreement. Consideration of similar breakdowns for several other targets reveals that these proportions alter with change in pattern, though, with one exception, not dramatically. The exception, however, is peculiarly interesting. The *circle* has a very high 'unitary' disappearance value, and in the best Gestalt tradition vanishes as a whole more than twice as often as any other figure tested. Table 3 shows the figures for unitary disappearance for a circle when compared with an ellipse of equal area and angular dimension.

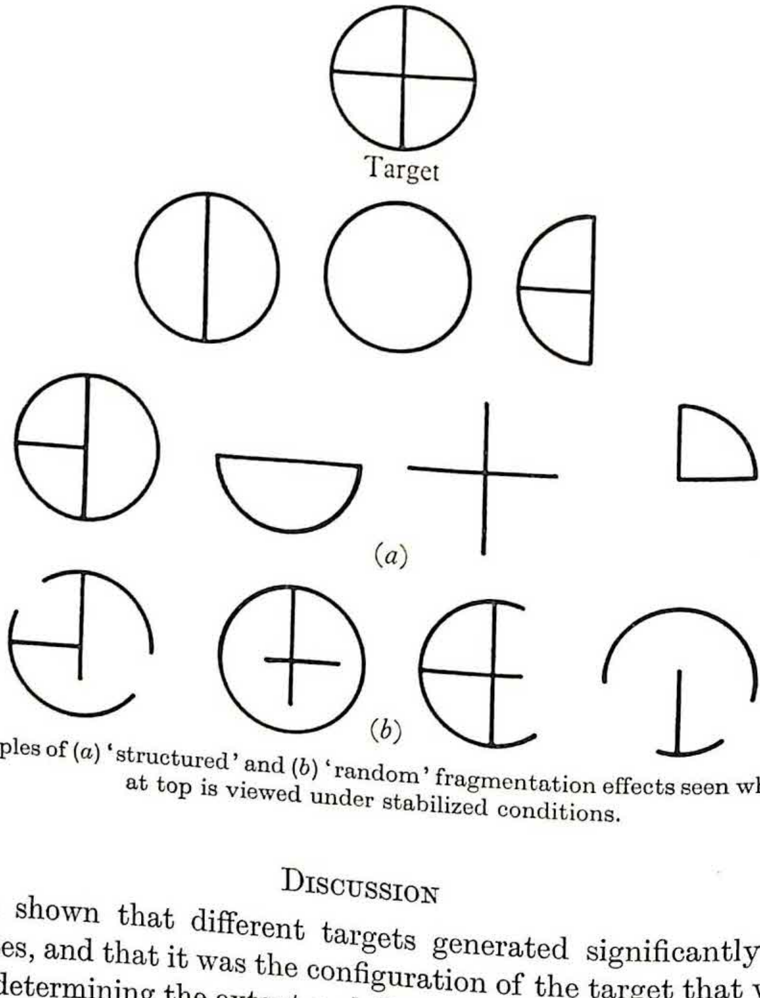


Fig. 8. Examples of (*a*) 'structured' and (*b*) 'random' fragmentation effects seen when the target at top is viewed under stabilized conditions.

Discussion

It has been shown that different targets generated significantly different disappearance rates, and that it was the configuration of the target that was of primary importance in determining the extent and direction of these differences. The proposition that disappearances of objects in the stabilized condition are not solely random or haphazard has also been examined, and, with one target at least, it has been shown that up to 20 % of the disappearances were of the type here called 'structured'. Though the work outlined must be counted as exploratory, and though it would be rash to build too much theory upon these results, it may be justifiable to ask whether the phenomena of disappearance can now be explained adequately by reference to retinal mechanisms only.

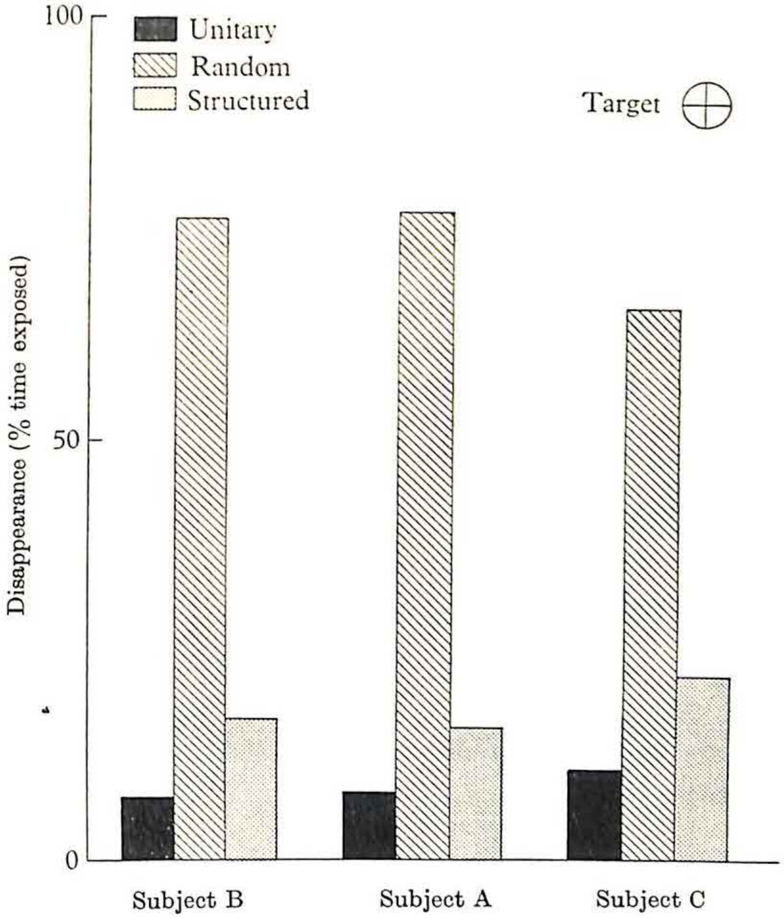


Fig. 9. Breakdown of percentage disappearance into 'unitary', random' and 'structured' disappearances (see text) for subjects A, B and C.

Table 3. *Comparison of 'unitary' percentage disappearances for circle and ellipse*

(The difference between the means is significant ($P < 0{\cdot}001$), showing that the circle tended much more to vanish as a unit.)

	Circle	Ellipse
	40·3	8·9
	33·8	17·7
	53·1	8·3
	44·9	26·7
	32·2	18·9
	34·4	6·6
	31·3	13·7
	19·5	8·9
Mean	36·1	13·7

It is, perhaps, not surprising that visual images disappear when good compensation is made for eye-movements. If the natural movements of the eye are inhibited or controlled in some way, we might well expect unusual events to occur in the visual system. The nervous system is undoubtedly geared to a changing environment, and most receptive areas of the body are eminently suited to detecting changes. More-

over, receptive cells require time to recover from heavy stimulation and may become fatigued or satiated if stimulation is prolonged. A first hypothesis is therefore the possibility that the disappearance effects reflect nothing more unusual than the satiation of retinal cells.

If, however, satiation were the *sole* cause of disappearance effects, one would expect that they would either be: (*a*) quite random (i.e. very few or no *structured* effects), or (*b*) quite unitary (all parts fading at once), or (*c*) if fragmentary, related in some simple and specific way to retinal locus. With regard to (*a*) it has been shown that, at most, 75 % of the effects were random and, with regard to (*b*), that less than 10 % were unitary, with the striking exception of the circle. As for (*c*), if retinal locus were vital, owing for instance to differential densities of receptors in the fovea and its periphery, then one might expect some degree of fragmentation of disappearance. With a figure such as the circle with inscribed cross one would expect frequent disappearances of the circle, perhaps with some sort of stunted cross remaining in the centre. But, as all subjects reported, much the most common effect was that one vertical or horizontal bar of the cross disappeared *as an entity*, leaving the circle and the remaining bar untouched.

To attribute the latter type of disappearance to satiation of retinal elements alone is difficult, and when one considers more complex configurations it becomes impossible, e.g. when one quadrant only of the circle–cross figure vanishes.

While it has for some time been agreed that the retina is something much more than a passive collector of 'yes–no' visual signals, there has remained a reasonable disinclination to consider it capable of making a differential response to different shapes. However, recent work such as that of Barlow & Hill (1962) clearly demonstrates that the rabbit retina contains ganglion cells capable of responding selectively to directional movement, and at least the first level of the receptive-field system uncovered by Hubel & Wiesel (1959, 1962) must lie in the retina for the units to be functional. Nevertheless, one should be chary of assigning too much work to the retina; if much recognition and perceptual classification processes lie in the retinal mosaic, then we may start asking ourselves why we have a cortex at all.

Two facts at least should be considered. (*i*) In our experiments we have noted the unitary action of lines subtending five and even ten degrees of arc. Since these are always foveally fixated the images of the lines must traverse both rod- and cone-free areas. It seems most unlikely that classificatory units of this size exist actually in the retina, knowing the anatomical variations which must occur over areas of this size. (*b*) Units such as those discovered by Hubel and Wiesel relate to the visual system of the cat. We might expect perceptual units mediating more complex forms to exist in the human visual system.

Finally, it is necessary to consider the possible role of contact lens slip in producing (*a*) partial fragmentation effects, and (*b*) the differential disappearance rates between targets of different shapes. Barlow (1963) has shown that contact lens slip with standard haptic lenses may be as great as 30 min. of arc. It is conceivable that with certain targets lens slip along the plane of a particular line might produce partial regeneration of an image, but such occurrences would seem likely to be rare. In fact, as shown above, structured configurations made up an appreciable percentage of the total effect; also, it has been shown elsewhere that similar fragmentation effects occur

under conditions of steady fixation alone, when no contact lenses are worn (Evans & Piggins, 1963), and when patterned targets are viewed as after-images (Bennet-Clark & Evans, 1963). As for (*b*), it might be supposed that the régime of eye-movements varied systematically according to the pattern presented. If this were so, one would certainly expect more and greater eye-movements with complex targets and hence more mechanical destabilization. But the present findings indicate that the more complex patterns induce the higher *disappearance values*, a fact suggesting that lens slippage of the kind investigated by Barlow is not an important variable in the studies described here.

The author wishes to express his thanks to Prof. M. D. Vernon and Prof. R. W. Ditchburn, F.R.S., for their advice and encouragement in the course of the experimental work reported in this paper. The author is particularly grateful to Miss D. F. Budge, who has patiently acted as subject in many of the experiments, and to Dr H. C. Bennet-Clark, Dr P. W. Nye and Mr D. C. West for their services in the same role. This investigation was supported in whole by Research Grant NB-01233 from the National Institute of Neurological Diseases and Blindness, Public Health Service, U.S.A.

References

Barlow, H. B. & Hill, R. M. (1963). Selected sensitivity to direction of motion in ganglion cells of the rabbit retina. *Science*, **139**, 412–4.

Barlow, H. B. (1963). Slippage of contact lenses and other artefacts in relation to fading and regeneration of supposedly stable retinal images. *Quart. J. exp. Psychol.* **15**, 36–51.

Bennet-Clark, H. C. & Evans, C. R. (1963). Fragmentation of patterned targets when viewed as prolonged after-images. *Nature, Lond.*, **199**, 1215–16.

Clowes, M. B. (1959). Eye movements and discrimination of brightness and colour. Ph.D. Thesis, University of Reading.

Ditchburn, R. W. & Pritchard, R. M. (1956). Stabilized interference fringes on the retina. *Nature, Lond.*, **177**, 434.

Evans, C. R. & Piggins, D. J. (1963). A comparison of the behaviour of geometrical shapes when viewed under conditions of steady fixation and, with apparatus for producing a stabilized retinal image. *Brit. J. physiol. Optics*, **20**, 1–13.

Fender, D. H. (1956). The function of eye movements in the visual process. Ph.D. Thesis, University of Reading.

Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–91.

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–54.

Pritchard, R. M. (1958). Studies of visual perception with a stabilized retinal image. Ph.D. Thesis, University of Reading.

Pritchard, R. M., Heron, W. & Hebb, D. O. (1960). Visual perception approached by the method of stabilized images. *Canad. J. Psychol.* **14**, 67–77.

(*Manuscript received* 14 *March* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 135–141
Printed in Great Britain

THE PERCEPTION OF ILLUSIONS AS A CONSTANCY PHENOMENON

By L. B. BROWN* and L. HOUSSIADAS†
University of Adelaide

The view, recently revived, that the non-veridical perception of visual illusions results from size constancy processes set up by the depth features of flat figures has been tested with a series of eight stimuli. Data collected from experiments with 44 subjects suggest that this view applies only to a limited range of visual illusions and that, therefore, it has not the generality claimed by its proponents.

Various theories have been put forward since the end of the last century to account for optical illusions. Some of them have stressed physiological factors at the retinal level, others, including the Gestalt theorists, have attributed illusions to the interplay of 'field forces', while a third group has emphasized factors of perspective, such as phenomenal depth, in accounting for illusory perception (see Kristof, 1961, for references to the early studies). Recently a number of writers, such as Tausch (1954, 1955), Kristof (1961), and Gregory (1962, 1963) have suggested extensions or elaborations of this third group of explanations. Although there are some differences between them, all three of these authors agree that the perception of optical illusions is a function of constancy processes.

The experiments to be reported in this study were designed to test the generality of this 'constancy' interpretation, with special reference to Gregory's (1963) suggestion that 'illusion figures may be thought of as flat projections of typical views of objects lying in three-dimensional space....The parts of the figures corresponding to distant objects are expanded and the parts corresponding to nearer objects are reduced' (p. 678). These distortions, according to Gregory, result when the constancy scaling process is set inappropriately to the distance of an observed object. Processes giving rise to such systematic distortions are not unknown in psychology. After-images, for instance, arising from the same retinal image look larger when projected on a distant screen and smaller when projected on a nearer screen (Emmert's Law). Gregory illustrates his interpretation by referring to a series of illusions drawn on normally textured paper, including those of Müller-Lyer, Ponzo and Hering, and by presenting two-dimensional, luminously painted wire models of some illusions in the dark, when they are reported to appear three-dimensional.

The following predictions, derived from Gregory's general account of visual illusions, were therefore tested on a series of eight illusory figures: (*a*) that the perceived distance between the illusory lines in halves of Wundt's and Hering's figures should be *larger* in the part of the background which appears to be farther away, or *shorter* in the part of the background which looks nearer; (*b*) that the perceived length of the horizontal lines in Ponzo's figure should be a function of apparent depth, with that part of the figure which looks further away having the longer horizontal line; (*c*) that

* Now at Massey University of Manawatu, New Zealand.
† Now at University of Thessaloniki, Greece.

when the background sets up a 'constancy scaling', as in Gibson's textured fields, parallel lines drawn on this background should appear to diverge in that part of the figure which looks farther away. The stimulus figures were presented on textured paper and therefore give what Gregory (1964, p. 303) would presumably regard as a 'rather indirect' testing of his form of the theory.

Experiment I

Method

The preceding predictions were tested on eight stimuli, shown in Figs. 1–6, each with a background giving rise to a process of 'constancy scaling'. The first two *stimuli* were the top and bottom halves of Wundt's figure and the third and fourth were versions of Hering's illusion; Figs. 1 and 2 were stimuli $1a$ and $2a$ when presented as shown, but when turned upside down they were stimuli $1b$ and $2b$, respectively. Stimulus 3 is shown as Fig. 3 and consisted of two parallel lines of equal length superimposed on one of Gibson's textured fields; stimulus 4 (Fig. 4) was Ponzo's illusion turned through 180°; and the last two stimuli (Figs. 5, 6) had vertical or horizontal parallel lines of equal length drawn on another of Gibson's textured fields. The stimuli were photographed to the same size (8 × 7 cm) and were presented in a different random sequence to each of twenty-six second-year students of psychology.

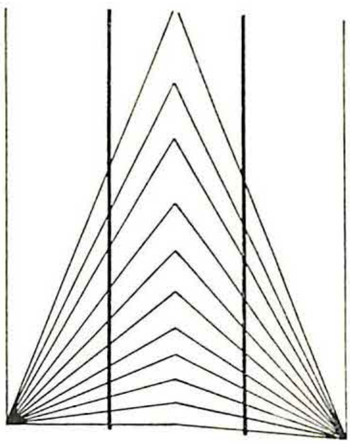


Fig. 1

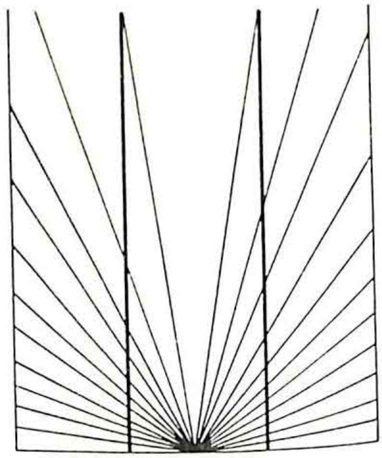


Fig. 2

Fig. 1. Half of Wundt's figure, used as stimulus $1a$ in this orientation and as stimulus $1b$ when turned upside down.

Fig. 2. Half of Hering's illusion, used as stimulus $2a$ in this orientation and as stimulus $2b$ when turned upside down.

For each figure, each subject was asked first to compare the illusory lines with standards which were shown separately. For stimuli $1a$, $1b$, $2a$, $2b$, 3, and 5 the first standard (A) showed two parallel vertical lines of equal length; in the second (B) these lines were diverging from the subject; while in the third (C) they converged towards the subject. Thickness, length and distance between the lines were similar to those of the stimulus figures. For stimuli 4 and 6, the first standard for comparison consisted of two horizontal parallel lines of equal length, the second and third standards had the top line shorter in the first case, and longer in the second, than the bottom line. Thickness, length and distance between the parallel lines were again similar to those of the stimulus figures. Subjects were next asked, 'In which direction does the illusory figure seem to lie? That is, if this were a three-dimensional picture, which part looks as if it would be nearer?' Finally the subjects were asked to 'look again at these illusory lines and compare them with the standards A, B, and C'. (Those subjects who changed their judgements between the first and third phase have been placed in a 'doubtful' category in the results.)

The stimuli were shown to the subjects at a distance of about 45 cm, in the frontal plane, and without any time limitations. Fixation areas were not stipulated but subjects were discouraged from fixating along the illusory lines, as this tended either to decrease or to eliminate the illusory effects of the figures. All subjects were naïve in that they had not participated in similar experiments previously.

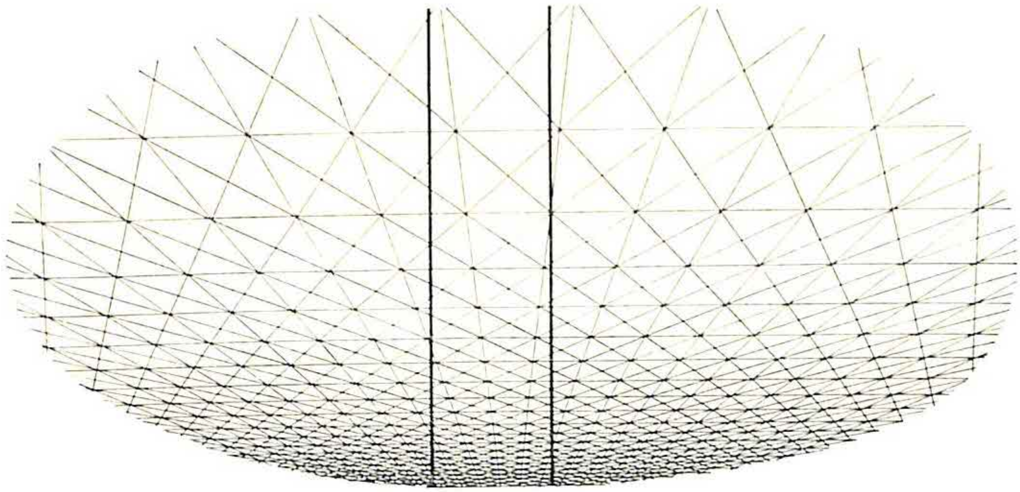


Fig. 3. Stimulus 3, two equal, parallel lines on one of Gibson's (1950) textured fields.

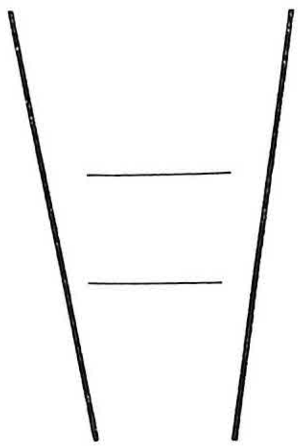


Fig. 4

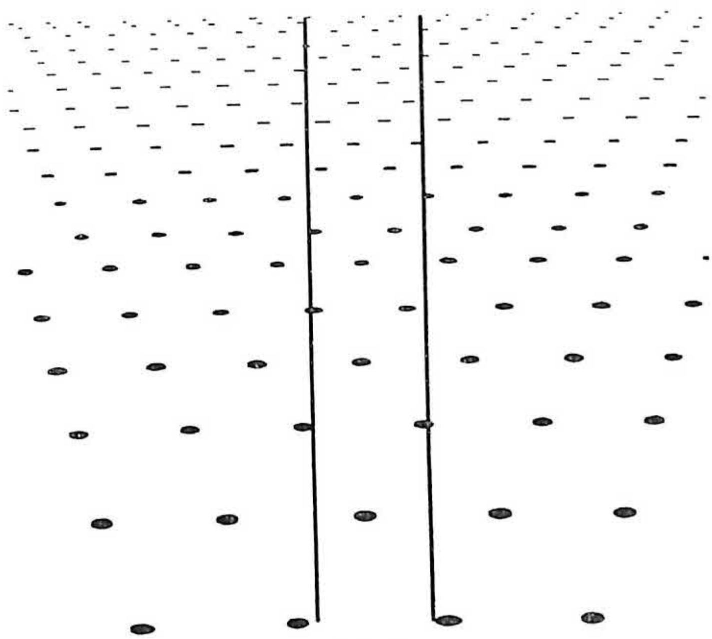


Fig. 5

Fig. 4. Stimulus 4, the Ponzo figure.

Fig. 5. Stimulus 5, vertical parallel lines on one of Gibson's (1950) textured fields.

Results

The data have been analysed in terms of the prediction that in general the illusory lines should converge, or in Ponzo's figure look shorter, at the nearer part of the background. The results are set out in Table 1. Judgements that were 'not as

predicted' are either 'illusory', when the subjects saw the illusion but did not report the predicted apparent distance, or 'veridical', when no illusion was reported but the background was seen as either nearer or further away. The 'doubtful' category covers judgements of distance in the background yet with both illusory *and* veridical responses given concerning the figured lines.

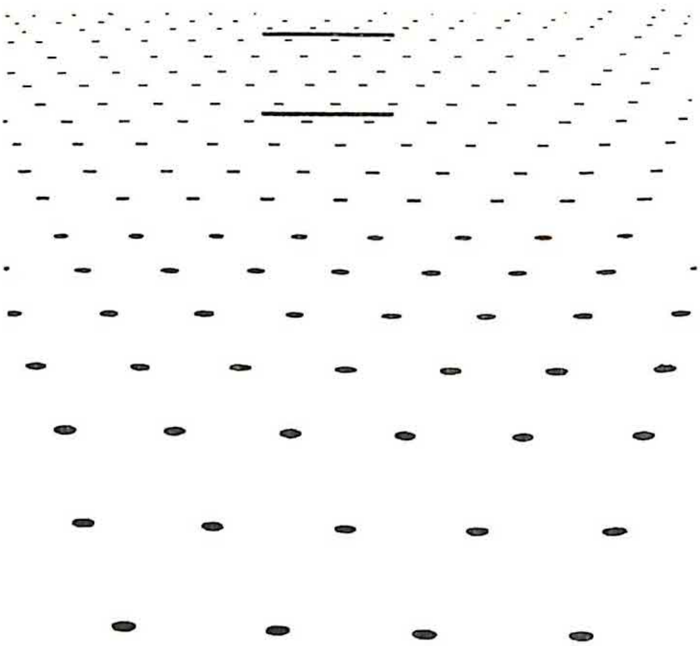


Fig. 6. Stimulus 6, horizontal parallel lines on the same field as in Fig. 5.

Table 1. *Frequencies of responses conforming and not conforming to Gregory's predictions* ($n = 26$)

Stimulus	Gregory's prediction	Not as predicted: Illusory	Not as predicted: Veridical	Doubtful
1*a*	9	9	5	3
1*b*	4	13	5	4
2*a*	14	10	2	0
2*b*	20	5	1	0
3	6	0	16	4
4	3	7	15	1
5	1	0	25	0
6	1	1	24	0

These results do not give unequivocal support to Gregory's view, and broadly confirm the present authors' statement advanced in a preliminary note (Brown & Houssiadas, 1964). Stimuli 2*a* and 2*b* give results in line with the predictions from Gregory, and stimulus 3 gives limited support in that in these figures convergence and divergence of the illusory lines seems to be associated with the parts of the figures that are respectively 'farther away' or 'nearer'. However, stimuli 1*b* and 4 do not confirm Gregory's interpretation as *divergence* is associated with the 'farther away' and convergence with the 'nearer' part of the background. The results from the rest of the stimuli are somewhat ambiguous. Although stimuli 5 and 6 have clear depth cues the superimposed lines were seen veridically in both cases. It is true that with

these two stimuli (particularly stimulus 5), some subjects reported that they saw the lines as perpendicular to the 'floor', but they could also see them as 'lying on the floor' with veridical perception of the 'illusory' lines. We did not collect any systematic data concerning this phenomenon, but it emphasises the possible independence of the background and the 'illusory' lines.

Experiment II

Method

The second experiment was carried out as a further test of Gregory's interpretation and, in fact, is a logical extension of the first experiment in which a number of the subjects reported spontaneously that they could see Wundt's and Hering's figures reversing. If such a change occurs in the background of the figures, it should, following Gregory, be associated with a change in the perception of the illusory lines. Thus, converging lines should appear to diverge, when the 'nearer' part of the figure is seen as 'farther away'.

The stimuli 1*a*, 1*b*, 2*a* and 2*b* from the first experiment were therefore presented in this order to a further eighteen subjects. These subjects were asked to compare the illusory lines with the same standards as before and they were then asked whether they could see the background as having depth and, if so, whether they could reverse it. As it is difficult to make these reversals, each subject was first shown drawings of a reversible book and the Necker cube in Solley & Murphy (1960, p. 276). Those who reported that they could reverse the background of the figures were asked to notice whether any change took place in the illusory lines as the apparent direction of the background changed.

Results

Table 2 shows the results of this experiment, and it is clear that reversing the background is not followed by any change in the appearance of the illusory lines. This suggests that the expansion or contraction of the illusory lines is not a function of apparent distance.

Table 2. *Frequencies of responses with reversible background* ($n = 18$)

Stimulus	Third-dimensional response	Reversed background	Illusion: Changed	Illusion: Not changed
1*a*	18	16	0	16
1*b*	17	10	0	10
2*a*	16	9	0	9
2*b*	18	2	0	2

Discussion

Views similar to Gregory's about a 'constancy scaling' process in visual illusions seem to have been suggested by Gibson (1950) when he incorporated illusory perceptions into the phenomena of size constancy by suggesting that the perception of size is a by-product of a constant scale of the visual world at various distances. Although Gregory (1964, p. 302) argues that Gibson's view of constancy 'precludes this kind of theory', Gibson (1950) suggested that this scale is what remains constant in perception, with the size of any object 'given by the scale of the background *at the point to which it is attached*, and that is why apparent size is linked to its apparent distance' (p. 181, italics in original). It is true that Gibson did not develop a theory of illusions, but his interpretation is derived from those illusions in which real objects

are represented in pictures, with three-dimensional forms deliberately drawn in two-dimensions. Such pictures, according to him, are not illusory in so far as they are substitutes for objects. They become illusory when they are objects in themselves (for example, black lines on white paper). It is not clear, however, if Gibson would consider that these illusions are produced by the same processes as are conventional geometrical illusions. He reported that, because of their surface texture, figures which normally give the impression of being compressed appear as pure depthless shapes lying in the frontal plane when textural cues are eliminated by darkness (Gibson, 1950, pp. 173–4). The fact that the cues which give rise to three-dimensional impressions in two-dimensional pictures do not operate when textural factors are absent may suggest the operation of different processes under different conditions.

Therefore some illusions may be thought of as pictures substituted for objects, with the illusory effects a function of distance cues in the background as in Gibson's cylinders (Gibson, 1950, p. 182). The same may apply to an intermediate range of stimuli, including Ponzo's figure, Gibson's textured field and the Müller–Lyer illusion, all cited by Gregory, in which there can be a correspondence with real objects.

There is however a series of geometrical illusions, including versions of Zöllner's illusion, the Wundt figure which we used as stimulus 2 and others cited by Houssiadas & Brown (1963, particularly Figs. 2–5 and 9), in which apparent distance cannot explain the illusory effects, partly because of a separation of figure from ground which does not allow the figure to fit into the ground. In these figures it is probable that parts, or 'small regions' (Marshall & di Lollo, 1963), are responsible for the illusory effects, and consideration needs to be directed to such features of these figures as the areas of fixation, the role of points of intersection of the lines, and the influence of the angle of inspection.

Therefore, Gregory's contention that 'the scaling can be set directly by depth features of flat figures' does not seem to be supported experimentally in all geometrical illusions. Not only do our results force this conclusion, but Green & Hoyle (1963) found the same to be the case with Poggendorff's illusion when it was presented under the reduced cue conditions used by Gregory. Balser (1963) also reports control experiments with Kristof's illusory figures in which the apparent length of two parallel lines of equal length drawn on the photograph of a landscape, and the apparent size of two dots of equal size on the Necker cube (one on the 'back' and the other on the 'front'), do not change when the figure is turned through 180° and the background becomes the foreground and vice versa. Gregory's point that the perspective does not change when figures are rotated through 180° is arguable (Gregory, 1964, p. 303). If perspective is 'perception of the relative distance of various objects...by means of the apparent size of the objects, relative size of parts, shading, etc.' (Warren, 1934), then clearly with 'forced' or 'spontaneous' reversibility, where the same part of the figure changes from 'nearer' to 'farther away', and vice versa, a change of perspective is involved. It is significant that, in the second experiment, changes in perspective in this sense were not followed by changes in the perception of the illusory lines. Similarly, in the first experiment, with Ponzo's figure rotated through 180°, seven of the ten subjects who saw the illusion did not associate distance with the length of the horizontal lines, as predicted from Gregory. It seems,

in other words, that the same cues do not necessarily give the same information about distance, and that if appreciation of distance changes then perspective in this sense changes. Gregory's studies of apparent depth in luminous illusions when extended to include reversible backgrounds may clarify these processes (cf. Gregory, 1964, p. 303).

We conclude that Balser's results, together with those of our experiments, show that 'direct' inquiry does not support the generality of a 'constancy scaling' account of visual illusions. The argument that illusions are three-dimensional interpretations of two-dimensional displays seems to apply only to those special figures that are intermediate between geometrical illusions where there is little consistent three-dimensional effect and those 'pictures' which are deliberately two-dimensional representations of objects in three dimensions.

References

Balser, M. (1963). Untersuchungen zu Kristof's Zweikomponententheorie der geometrisch-optischen Täuschungen. *Arch. ges. Psychol.* **115**, 307–29.

Brown, L. B. & Houssiadas, L. (1964). Illusory perception as a constancy phenomenon. *Nature, Lond.*, **204**, 302.

Gibson, J. J. (1950). *The Perception of the Visual World.* Boston: Houghton Mifflin.

Green, R. T. & Hoyle, E. M. (1963). The Poggendorff illusion as a constancy phenomenon, *Nature, Lond.*, **200**, 611–2.

Gregory, R. (1962). How the eyes deceive. *The Listener*, no. 1736, pp. 15–16.

Gregory, R. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature, Lond.*, **199**, 678–80.

Gregory, R. (1964). A reply. *Nature, Lond.*, **204**, 302–3.

Houssiadas, L. & Brown, L. B. (1963). The effect of viewing slant and distance on some visual illusions. *Aust. J. Psychol.* **15**, 108–12.

Kristof, W. (1961). Über die Einordnung geometrisch-optischer Täuschungen in die Gesetzmässigkeiten der visuellen Wahrnehmung, *Archiv. ges. Psychol.* **113**, 1–48, 127–50.

Marshall, A. J. & Di Lollo, V. (1963). Hering's illusion with improvement of the stimulus in scotopic and photopic vision. *Amer. J. Psychol.* **76**, 644–52.

Solley, C. M. & Murphy, G. (1960). *Development of the Perceptual World.* New York: Basic Books.

Tausch, R. (1954). Optische Täuschungen als artifizielle Effekte der Gestaltungsprozesse von Grössen und Formenkonstanz in der natürlichen Raumwahrnehmung. *Psychol. Forsch.* **24**, 299–348.

Tausch, R. (1955). Nichtbewusste (sog. unbewusste) Vorgänge bei der optischen Grössenwahrnehmung von Gegenständen. *Psychol. Forsch.* **25**, 28–64.

Warren, H. C. (1934). *Dictionary of Psychology.* London: Allen and Unwin.

(*Manuscript received* 6 *November* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 143–146
Printed in Great Britain

MATCHING LOUDNESS AND VOCAL LEVEL: AN EXPERIMENT REQUIRING NO APPARATUS

BY R. J. IRWIN
University of Auckland
AND A. WILLIAM MILLS
Tufts University

Both the loudness of a sound and the apparent magnitude of a self-produced vocal response (autophonic response) can be described as power functions of sound pressure, although with different exponents. Two scales, one of loudness, the other of autophonic level, can therefore be specified as a function of the same sound pressures. Under these circumstances theory predicts that loudness should be proportional to the square root of apparent autophonic level. Two separate but similar classroom experiments were performed, in each of which one person made autophonic responses of six different magnitudes whose loudnesses were judged by the members of the class. The obtained exponents between the two scales so erected was 0·52 in one demonstration and 0·53 in the other, thus confirming the prediction. The experiments used no apparatus since the only specification of the stimulus required was that the sound pressure produced by one subject should be the same as that judged by another.

INTRODUCTION

Experiments on cross-modal matching (Stevens, 1959) require that subjects judge when two stimuli, sensed by different modalities, are equal in apparent magnitude. By this means pairs of stimuli of apparently equal magnitudes are determined. If the subjective magnitude of each stimulus bears a known relation (as specified by some psychophysical function) to its physical magnitude then it follows that pairs of stimuli corresponding to apparently equal magnitudes must bear a predictable relation to each other. For example, if the sensation of two different stimulus-dimensions are power functions of their respective stimuli, then it can be shown that such stimuli, matched in subjective magnitude, should also be related by a power function, but with an exponent equal to the ratio of the exponents of each dimension. This deduction, which is due to Stevens, therefore provides a check on the validity of the separate scales of apparent magnitude—a check which is, however, a necessary consequence but not a sufficient proof of their validity, because other functions (such as a logarithmic one) lead to the same deduction.

A special case of cross-modal matching has been reported by Lane, Catania & Stevens (1961). They asked subjects to match the apparent magnitude of the vocal response (autophonic response, as the authors called it) to the loudness of a white noise. From independent experiments (e.g. Stevens, 1955) the loudness of white noise, when judged by one of the direct methods such as the method of magnitude estimation, is proportional to the 0·6 power of the sound pressure. The apparent level of a subject's autophonic response, as measured by the method of production, is proportional to about the 1·17 power of the sound pressure produced (Lane *et al.* 1961). It follows that if a subject adjusts the loudness of a noise to seem similar in magnitude to one which he himself is producing, the sound pressure of the vocal response should be the 0·6/1·17 power of the sound pressure of the equated noise.

This prediction was confirmed by Lane *et al.* for the autophonic response of the vowel [ɑ], and the loudness of white noise.

Lane's experiment is a special case of cross-modal matching because the two modalities (speech and hearing) are both described as functions of the same physical dimension, namely, sound pressure. For this reason it affords an opportunity of performing the converse experiment: instead of measuring pairs of stimuli corresponding to equal sensations, one can measure pairs of sensations corresponding to equal stimuli. The magnitude of the stimuli need not be known, so long as for any pair of judgements they are the same.

Thus given the following psychophysical functions (and ignoring constants which govern the size of units),

$$L = P^{0\cdot 6} \quad \text{and} \quad A = P^{1\cdot 17}, \tag{1}$$

where L is loudness, A is apparent magnitude of autophonic response, and P is sound pressure; then, if L and A are found for identical values of P,

$$L^{1\cdot 17} = A^{0\cdot 6} \tag{2}$$

or

$$\log L = 0\cdot 6/1\cdot 17 \log A\,; \tag{3}$$

$\log L$ should then be a linear function of $\log A$, with a constant of proportionality of 0·51. Note that physical magnitudes do not appear in equation (3).

This function provides the basis for an instructive demonstration of cross-modal matching that can be performed in the classroom. Two such demonstrations are now described.

METHOD

Each author carried out an independent but procedurally similar experiment. In one experiment the subjects were nineteen students of a second-year psychology class at the University of Auckland, and in the other they were twenty-five students of a first-year course at Tufts University. The subjects were read instructions which asked them to estimate the loudness of the vowel [ɑ] which would be said by the teacher. The first vowel emitted was to be the standard stimulus, and was assigned a loudness of ten; other vocal productions were to be given numerical estimates proportional to their loudnesses, bearing in mind that the first one had a loudness of ten. (This is the method of magnitude estimation.) The teacher then produced six different autophonic levels whose magnitudes bore a specified relation to the magnitude of the standard: these relations were one-fifth, one-half, once, twice, three times, and ten times. The standard was presented first, and each of the five remaining magnitudes was presented twice in a predetermined and random order. (This is the method of magnitude production.) The modulus of the autophonic scale was set by producing the standard at what was deemed to be a conversational level.

RESULTS

The results of the two experiments are shown separately in Fig. 1. The experimental points represent the median estimations of the loudness of each of the vocal levels produced. The filled circles are the medians of 38 judgements from 19 subjects, and the open circles are those of 50 judgements from 25 subjects. Each point represents two sets of judgements described as a function of the same stimuli: one set of judgements about loudness, and the other about autophonic level. The triangle at the co-ordinates (10, 10) represents the standards which were arbitrarily assigned to that position on both scales. The vertical lines around each point extend plus and minus one quartile from the medians.

The solid line is the theoretical outcome predicted from the separate psychophysical functions of equation (1). The constant 10 in the equation that describes the line merely reflects the choice of units. The method of least squares was used to determine what power function fitted the medians best. The best-fitting exponent for the filled circles is 0·52, and that for the open circles is 0·53; the theoretical equation thus represents the data well. The data may therefore be described by saying that the judgements of loudness were approximately proportional to the square-root of the apparent magnitude of the autophonic response. The largest discrepancy between theory and experiment is shown by the most right-hand points, where one of the obtained medians underestimates the theoretical outcome of 31·62 by 6·62 units, and the other overestimates it by 8·38 units.

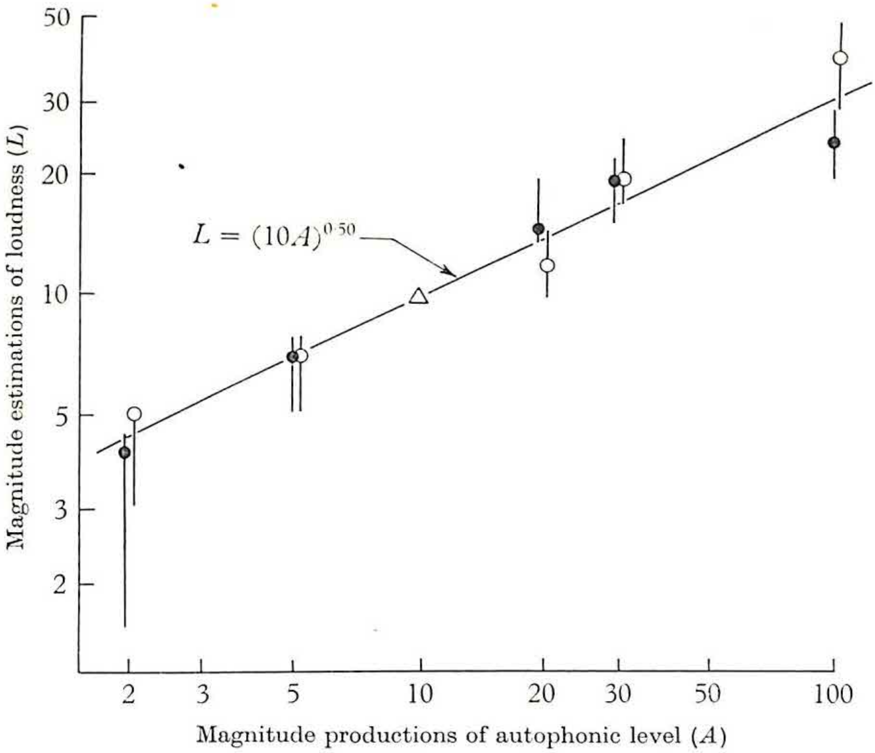


Fig. 1. Estimates of the loudness of a vocal response plotted as a function of its apparent autophonic level. The filled and open circles are independent estimates from different experiments. The filled circles have been displaced slightly to the left of the open circles for the sake of clarity. The triangle represents the standards which were arbitrarily assigned to that point on both scales. The vertical lines extend ± one quartile from the medians. The straight line is the theoretical outcome predicted from the separate psychophysical equations which specify each subjective scale as a function of sound pressure.

Discussion

Three characteristics distinguish these experiments from the usual one on cross-modal matching. First, instead of determining the locus of stimuli corresponding to equal sensory magnitudes, they determined the locus of sensory magnitudes corresponding to equal stimuli. Secondly, whereas the same subject usually makes the judgements of both stimuli, here the person who produced the autophonic responses was not the one who estimated their loudness, or, at any rate, their loudness un-

confounded by the effects of producing them. Thirdly, both sensory modalities were scaled on the same physical dimension, not different ones. But since the sense of hearing is the only one capable of transducing small sound pressures, the psychophysical equation for the autophonic response may be somewhat misleading. Its exponent cannot reflect, as is claimed for other exponents (Stevens, 1961), the operating characteristic of its transducer, for the ear is the transducer of sound pressure, and its exponent is 0·6. Although it may be possible to specify both loudness and autophonic level as a function of sound pressure, the judgements must be made of different stimuli, otherwise loudness could not be proportional to the square root of autophonic level when both are described as functions of the same stimuli. Warren (1962), on the other hand, suggested that judgements of both loudness and vocal level are based on estimates of distance: a sound twice as loud as another has a sound pressure of one twice as near, and a vocal level twice another is one whose sound pressure is projected twice as far. This explanation must therefore invoke additional hypotheses to account for the non-linear relation between loudness and vocal level reported by Lane *et al.* (1961), and also found in the present experiments.

Despite these differences from the customary experiment on cross-modal matching, or perhaps because of them, we believe these experiments illustrate a feasible and instructive demonstration which can be performed in the classroom without any apparatus.

REFERENCES

LANE, H. L., CATANIA, A. C. & STEVENS, S. S. (1961). Voice level: autophonic scale, perceived loudness, and effects of sidetone. *J. acoust. Soc. Amer.* **33**, 160–7.

STEVENS, S. S. (1955). The measurement of loudness. *J. acoust. Soc. Amer.* **27**, 815–9.

STEVENS, S. S. (1959). Cross-modality validation of subjective scales of loudness, vibration, and electric shock. *J. exp. Psychol.* **57**, 201–9.

STEVENS, S. S. (1961). Psychophysics of sensory function. In *Sensory communication* (ed. W. A. Rosenblith). New York: M.I.T. and Wiley.

WARREN, R. M. (1962). Are 'autophonic' judgements based on loudness? *Amer. J. Psychol.* **75**, 452–6.

(*Manuscript received* 1 *September* 1964)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 147–156
Printed in Great Britain

REPETITION AND VERBAL LEARNING

By PETER B. WARR
Department of Psychology, University of Sheffield

An analysis is developed of the possible consequences of presenting for learning an item amongst others in a list, and a hypothesis is generated about the factors determining which of these outcomes will occur. A technique for measuring the effect of repetition is described and discussed. Two experiments are reported which test predictions about the manner in which task difficulty mediates the effect of repeated presentation. The results appear to confirm the predictions. Some implications of these experiments are considered.

Until recently it was customary to assume that all learning proceeded incrementally. It was taken for granted that repeated presentation of an item led to successive increments in associative strength. A number of writers have recently questioned this assumption and have produced evidence of what appears to be all-or-none learning. These studies have been reviewed and evaluated by Postman (1963) and Kay (1965).

It is now usual to argue in favour of *either* the incremental hypothesis *or* the all-or-none hypothesis, as if only one of these could be valid. But it is more likely that learning in some conditions is incremental and in other conditions is all-or-none. It is intended in this paper to expand this suggestion and to report experiments that test predictions about the variables mediating the effect of repeated presentation. The possible effects of repetition of items must first be analysed. This analysis will be presented in terms of a paired associate learning task.

It is clear that, if all-or-none learning is to occur, subjects must in some sense fail to process some of the information presented to them. It is therefore postulated that subjects may 'select' certain items from the list presented. Selection of an item is taken to imply that information about the item is processed by the subject; and conversely information about a non-selected item is not processed by the subject. It is probable that there are differences in the degree to which items are selected in this way. Selection may be an intentional or an unintentional activity. Examples of intentional selection are provided by subjects experienced in laboratory learning tasks who often deliberately adopt the strategy of paying attention to part of the list to the exclusion of the remaining items. Unintentional selection is exemplified by the subject who particularly attends to an item which includes the initials of his name and who thereby fails to process information about other items. Among the variables which may determine subjects' selection of items are the 'meaningfulness' of certain stimulus and response terms or of the linking of two terms of a pair, inter-item similarity, time schedules, serial order effects, instructions given and subjects' conception of the task.

According to this usage of 'selection', learning can only take place after an item has been selected, but selection of an item does not always result in that item being learned. It follows from this that when items are presented for learning for the first time, three outcomes are possible.

(1) An item is not selected for learning. In this case no increment in associative strength has been generated. The item will not be correctly recalled.

(2) An item is selected for learning and is correctly recalled. In this case a fairly large increment has been generated.

(3) An item is selected for learning and is not correctly recalled. In this case some increment in associative strength has been generated.

The all-or-none and incremental hypotheses about the learning process may now be described in terms of this analysis. The all-or-none hypothesis may be said to assume that on each trial items in a list fall either into category 1 or into category 2. This indicates either that an increment of sufficient magnitude to permit recall has been generated or that no increment in associative strength has followed the presentation. (The magnitude of those increments that have been generated can presumably vary within a limited range, so that there is a sense in which the expression '*all-or-none*' is misleading.) The incremental hypothesis may be said to assume that on each trial items in a list fall either into category 2 or into category 3. This indicates that an increment has been generated in the case of all items, and that the variation in magnitude of increments is fairly large. In terms of the present framework it will be said that the all-or-none factor is alone operative when all items in a list fall into categories 1 or 2. And it will be said that the incremental factor is alone operative when all items in a list fall into categories 2 or 3. It is now possible to describe an outcome which appears not to have been considered in earlier discussions of this topic—that items in a list fall either into category 1 or into category 2 or into category 3. We may say that when items fall into all the three categories both the all-or-none and the incremental factors are operative. This stage of the argument may be summarized in the form presented in Table 1.

Table 1. *Conditions under which the all-or-none and incremental factors will operate*

Category	Selected for learning	Increment in associative strength	Correctly recalled	Factors operating	
1	No	No	No	All-or-none (1, 2)	All-or-none and incremental (1, 2, 3)
2	Yes	Yes	Yes	Incremental (2, 3)	
3	Yes	Yes	No		

It is hypothesized that the proportion of items falling into each of the three categories varies between learning tasks and between conditions of the same task; the conditions in which learning occurs partly determine which factors are operating. In other words, it is postulated that certain situational variables mediate the effect of repeated presentation of items. It is assumed that the most important of these variables are concerned with the difficulty of the task.

At a relatively low level of task difficulty single selection learning of all selected items is assumed to be possible; in this case the all-or-none factor is alone operative. As the level of task difficulty is increased, subjects will no longer be able to recall correctly all items after a single selection, so that the incremental factor will also become operative. The incremental factor will be alone operative only when all items in a list are selected; it is likely that this will occur relatively infrequently in laboratory tasks. In brief then, the 'dual factor' hypothesis assumes that either one or both of the

factors described here may be operative in a particular task. It is hypothesized that certain situational variables determine whether the incremental factor, the all-or-none factor, or both factors are operative.

The task difficulty variables which are assumed to determine the effect of repetition include the time intervals and number of items employed. In addition it would appear that when subjects are required to make verbal responses during presentations of a list (i.e. to read aloud the stimulus and response terms), the level of task difficulty is increased; the introduction of verbal responses is of course analogous to a reduction in time intervals. Variations in task difficulty are assumed to be responsible for the seemingly inconsistent pattern of results of experiments in which unlearned items are replaced on each trial. In experiments of this design results apparently supporting the all-or-none hypothesis have been obtained by Clark, Lansford & Dallenbach (1960), Kristofferson (1961, Expt. 1), Reed & Riach (1960), Rock (1957) and Rock & Heimer (1959). On the other hand, a number of experiments of this kind have been reported in which the incremental factor appears to have been operative. It may be noted that the level of task difficulty in these experiments was markedly higher than in the studies referred to above. In the case of this second group of experiments (e.g. by Battig, 1962; Gregg, Chenzoff & Laughery, 1963; Postman, 1962; and Underwood, Rehula & Keppel, 1962), the task involved either verbal response during presentations, *or* relatively short time intervals, *or* both these conditions. The importance of these situational factors may be further emphasized by reference to experiments of this design by Kristofferson (1961) and Lockhead (1961). The results of Kristofferson's first study were of such a kind to suggest that only the all-or-none factor was operative. On the other hand, the introduction into his second experiment of verbal responses during presentations necessitated a significant amount of incremental learning. The same effect was noted by Lockhead (1961) when he cut presentation time from 8 to 2 sec. These changes in level of task difficulty are assumed to have led to changes in subjects' opportunity to select items from the list and to correctly associate selected items on the subsequent test.

If each item in a list is attended to briefly, all items may be said to have been selected. But it is unlikely that all items will be correctly recalled in these conditions. An experiment bearing directly upon this point has been reported by Brackett & Battig (1963). These investigators compared the performance of subjects who had received different pretraining treatments. Whereas one group was instructed during pretraining to learn only part of the list (in the present terminology, to select certain items), the other group was instructed to learn all the list (i.e. to select all items). The performance of the latter group in learning an additional list in which unlearned items were replaced was significantly worse than that of a control group. On the other hand, the group which had been trained to select a proportion of items was not worse than the control group. This may be interpreted to mean that the group which had been trained to select part of the list correctly associated all selected items after a single selection, whereas the other group selected all items but could not correctly associate all after a single selection. It might in this way be the case that *only* the incremental factor was operative for the latter group; this would be so if subjects followed instructions and selected all items. The all-or-none factor can only be operative if some items are unselected.

Predictions about task difficulty

An experimental procedure which generates a convenient index of which factors are operative in a learning situation has previously been reported (Warr, 1963). This procedure requires that items which are correctly recalled be replaced with randomly selected new items. In this way none of the items presented in any list has yet been correctly recalled. Although lists after the first contain items which have received varying numbers of presentations, the length of the list remains constant.

If items are categorized according to the number of presentations they have received, it becomes possible to compare the probability of correct recall, $p(C)$, for each category. The $p(C)$ measure is of the relative frequency kind, being the proportion of items which are correctly recalled after each number of presentations. For example, the probability of correct recall after two presentations is the proportion of those items presented twice which are correctly recalled after this number of presentations. The probability of correct recall after 1, 2, ..., n presentations may be abbreviated to $p(C_1)$, $p(C_2)$, ..., $p(C_n)$. The pattern of $p(C)$ values with increasing numbers of presentations indicates which factors are operative in a particular situation.

In conditions where the *all-or-none* factor alone is operative (i.e. where all items fall into categories 1 or 2) we expect $p(C)$ to remain constant despite increasing numbers of presentations. A straightforward test of this is provided by the comparison of $p(C_1)$ and $p(C_{2+})$. When the all-or-none factor alone is operative $p(C_{2+})$ will not differ significantly from $p(C_1)$.

When the *incremental factor* is operative, $p(C_{2+})$ will be significantly greater than $p(C_1)$. In addition, however, it is possible to distinguish between conditions in which the incremental factor alone is operative and those in which both factors are operative. This distinction may be made on the basis of the pattern of $p(C)$ values above $p(C_2)$. When the incremental factor alone is operative (i.e. when all items fall into categories 2 or 3), $p(C)$ will continue to increase with increasing number of presentations beyond the second.

On the other hand, when *both factors* are operative (i.e. when all items fall into categories 1, 2 or 3), it seems likely that $p(C)$ will not increase with increasing numbers of presentations beyond the second. This prediction derives from the assumption that an item selected but not correctly recalled on presentation n is not necessarily selected on presentation $(n+1)$ under these conditions If there is only a limited overlap between items selected on successive presentations and increments are assumed to decline with time, the proportion of items liable to be correctly recalled after more than one selection is likely to be independent of number of presentations. But $p(C_{2+})$ will be greater than $p(C_1)$ because of the benefit of the accruing to those items which are selected on occasions separated by only a short interval.

In accordance with the framework developed earlier it may be predicted that differences in task difficulty will be accompanied by differences in the pattern of $p(C)$ values. In relatively difficult tasks it is likely that $p(C_{2+})$ will be significantly greater than $p(C_1)$. But when the task is easier, $p(C_{2+})$ is not likely to exceed $p(C_1)$. Two experiments to test these predictions are reported below. The general approach is to require subjects to learn material of the same kind and under conditions similar

to those used in an earlier experiment (Warr, 1963). One aspect of task difficulty is varied in each of the experiments, whilst all other variables are held constant. The observed pattern of $p(C)$ values is compared to that found in the previously reported experiment. In this case the results indicated that both factors were operative in the paired associate task employed. The task was a relatively difficult one in that verbal responses were required during presentations which were made at a comparatively rapid rate.

Experiment I

Method

It was decided to relax the conditions obtaining during presentations of material in this experiment. It was predicted that in this case the incremental factor would no longer be required: $p(C_{2+})$ would not exceed $p(C_1)$. This argument was supported by the apparent lack of incremental learning in Rock's (1957) conditions, which constituted a very much easier task. A further experiment was therefore carried out in which procedure, conditions, materials, subject population, instructions, etc., were the same as previously, but in which Rock's time intervals were used. These are presented below for comparison.

	Rock's (1957) exp. (sec.)	Warr's (1963) exp. (sec.)
Presentation of each pair	3	3
Interval between pairs	5	0
Before recall test	Not stated	5
Test presentation of stimulus term	5	5
Interval between test items	0	0
After recall test	30	10

A 10 sec. interval before each recall test was employed, and (as in Rock's experiment) subjects were not required to make verbal responses during presentations. Apart from this manipulation of time intervals, conditions were those of the original experiment. The letter-number pairs which had previously been used were also employed in this experiment.

Table 2. *The relation between C-score and number of presentations in Expt. I*

	Successive presentations of an item					
	1	2	3	4	5	6
Items presented	1016	594	356	203	129	67
Total C-score	340	196	126	49	47	20
$p(C)$	0·335	0·330	0·354	0·241	0·364	0·299

Results

The relevant findings from Expt. I are shown in Table 2. The values of $p(C)$ in this Table are based upon group data, and since not all subjects contribute to the values for more than four presentations these should be considered unreliable. Comparison of $p(C_1)$ and $p(C_{2+})$ values for individual subjects indicates that the all-or-none factor is alone operative in these conditions (Wilcoxon $T = 195$, $n = 33$, $P > 0{\cdot}05$), whereas the incremental factor was present to a very significant degree in the conditions of the earlier experiment. This difference ensuing from the change in presentation conditions was shown to be significant by comparing the samples of difference-scores, $p(C_{2+})-p(C_1)$, from the two experiments. The effect of repetition was found to be significantly different in the two sets of conditions ($U = 219$, $P < 0{\cdot}01$).

Experiment II

The level of task difficulty is, however, determined not only by conditions obtaining during *presentation* of a list. Consider two groups of subjects for whom conditions during presentation of a list are identical. If the time available for recall of each item is 5 and 10 sec. respectively for the two groups, it is possible that the 10 sec. group would find the learning task easier. This would be because this group has longer to recall items which the former group might have 'learned' but are not able to recall in the 5 sec. available. In other words, if 5 sec. is not an adequate time for recall of all items selected, an extension of this recall interval should allow recall of a greater number of selected items. And if extension of the recall interval allows a greater number of selected items to be correctly associated, it increases the number of correct associations that will be made after a single selection. We might predict on this basis that a fairly large increase in the recall interval in a learning task will yield a situation in which incremental learning is unnecessary since all selected items can be correctly associated after a single selection.

Method

A further paired associate learning experiment of the same design was carried out to test this prediction. Conditions during presentations were the same as in the original experiment (Warr, 1963), but the time available for recall of items was increased from 5 to 10 sec. All other variables were held constant.

Table 3. *The relation between C-score and number of presentations in Expt. II*

	Successive presentations of an item					
	1	2	3	4	5	6
Items presented	585	430	290	198	130	95
Total C-score	120	108	67	38	20	17
$p(C)$	0·205	0·251	0·231	0·192	0·154	0·179

Results

The effect of repeated presentation of items upon the performance of the group is shown in Table 3. As usual, data for the larger numbers of presentations must be treated with caution. Analysis of the performance of individual subjects indicates that $p(C_{2+})$ is not significantly greater than $p(C_1)$ (Wilcoxon $T = 112$, $n = 24$, $P > 0{\cdot}05$). It appears that repetition of items does not lead to an increase in the probability of correct association in these conditions. The significance of the difference between the effect of repetition in this and in the original experiment was tested in the same way as for Expt. I. The difference was again significant ($U = 137$, $P < 0{\cdot}01$).

It is difficult to account for this finding except in terms of the hypothesis developed earlier. According to this it is assumed that a longer, 10 sec., recall interval allows all correct responses to be emitted after a single selection of the item, so that no incremental factor is operative. But when the recall interval is only 5 sec., as in the original experiment, single selection learning is not always possible; some items are correctly associated after a single selection whereas others require more than one selection before being correctly associated.

Further support for this hypothesis comes from the number of correct responses

emitted (C-score) on each test in the two experiments. Mean C-score per subject on each test in this experiment is 2·57, and mean C-score for the first six tests in the original experiment was 2·34. Although more items are correctly associated throughout Expt. II, this increase does not reach statistical significance. Predictions may be derived from the dual factor hypothesis about differences in C-score over the series of tests in the two experiments.

Conditions are identical on *Presentation* 1 of both experiments, but, since it is hypothesized that more items can be correctly associated after a single selection in the conditions of Expt. II, we predict that C-score for *Test* 1 will be at a significantly higher level in this experiment than in the original experiment. These values are 2·08 and 1·13 respectively for the two experiments so that this prediction is confirmed. Furthermore, according to the dual factor hypothesis items presented on each trial after the first in the original experiment are not selected, or are correctly associated after a single selection, or are correctly associated after more than one selection. On the other hand, items presented on trials after the first in Expt. II are all assumed either to be not selected, or to be correctly associated after a single selection. The operation of the incremental factor on tests after the first of the original experiment might therefore be supposed to result in a higher C-score for those tests. It has, however, been hypothesized that the conditions of the present experiment result in more items being correctly associated after a single selection than are so associated in the earlier experiment. This may be reformulated in the following way. The incremental factor necessarily contributes more to C-score on tests after the first in the original experiment, since this factor is not operative in Expt. II. On the other hand, the all-or-none factor contributes more to C-score on tests after the first in Expt. II, since more items selected for the first time can be correctly associated in the conditions of this experiment than in the conditions of the original experiment. Whereas the incremental factor is 'stronger' over trials after the first in the original experiment, the all-or-none factor is 'stronger' over the corresponding trials in Expt. II. We might expect, therefore, that C-score on comparable tests after the first will be approximately the same in both experiments. The mean values are 2·58 for the original experiment and 2·66 for the present one and can be seen to meet this expectation.

It is not easy to generate a different hypothesis to account for the observed pattern of C-score per test If it were the case that the subjects in Expt. II were faster learners than those in the earlier one (which would account for the C-score of 2·08 and 1·13 respectively on test 1), C-score for tests after the first should also be significantly higher in this experiment. This is not so. It may also be mentioned that the subjects were drawn from the same population for both experiments and are not expected to differ in learning ability.

But it is possible that the absence of incremental learning in the conditions of the present experiment is partly due to the increased interval between presentations resulting from a longer recall period. The mean interval between two successive presentations of an item in the original experiment was 111 sec., whereas in Expt. II the mean length of this interval was 171 sec. Consider a sample of items which are selected for learning in the original experiment. It is assumed that some increment in the associations between stimulus and response terms follows this selection.

Suppose, however, that this sample of items is not correctly associated on the following recall test. The mean interval before each item is presented again is 111 sec., and if an item is selected for a second time, we assume that a further increment is added. But it is reasonable to postulate some decline over time in the partial association formed after the first selection. It is possible that in a comparable sample of items in Expt. II this partial association has declined to a greater extent, since an average of 171 sec. has elapsed between presentations. In other words, it might be the case that (contrary to the explanation suggested previously) not all selected items in Expt. II are correctly associated after a single selection; but that no incremental learning takes place in these conditions because the increment in the strength of an association between terms of a selected item has declined to zero by the time a second selection of the item is made.

This postulated importance of inter-selection interval cannot wholly account for the absence of incremental learning in this experiment since it cannot explain the significantly greater C-score on Test 1. This seems to be adequately explained by the assumption that the longer recall interval permits more responses to be emitted correctly after a single selection. It may be the case, however, that the longer interval between presentations of an item is *partly* responsible for the constancy of the observed $p(C)$ values in a manner described in the previous paragraph. A combination of the two postulates seems to offer the most satisfying explanation of the absence of incremental learning in the conditions of Expt. II. Forgetting is clearly an important aspect of learning, but adequate techniques for studying the interrelationship between the two processes have yet to be developed.

Discussion

Insofar as recognition is generally regarded as being in some sense easier than recall (e.g. Achilles, 1920; Postman, Jenkins & Postman, 1948), we might predict that an easier replication of the original experiment (Warr, 1963), in which learning is assessed by recognition rather than by recall, will not involve the incremental factor. An experiment has been reported (Warr, 1964*b*) which differed from the original experiment only in that learning was assessed by recognition from a 5-item ensemble rather than by recall. In these conditions it was found that $p(C_{2+})$ did not differ significantly from $p(C_1)$.

Experiments have also been reported in which an attempt was made to study the response-learning and associative phases of paired-associate learning (Warr, 1964*a*). Response-learning was seen to occur on an all-or-none basis, whilst the associative phase involved both factors. It is now considered likely that these results reflect differences in the level of task difficulty, and that the factors operative in each phase are dependent upon changes in this variable. Empirical tests of this possibility may readily be made.

Other aspects of the hypothesis may be investigated experimentally. For instance, no task conditions have been employed in which the incremental factor is seen to be operative alone, although this may have been the case in Battig's (1962) experiment which used a design similar to the present one. For the incremental factor to be alone operative, all items in a list have to be selected; the experimenter may have to induce an appropriate set in the subjects before this takes place (e.g. Brackett & Battig,

1963). Differences in learning ability also constitute a promising line of investigation; it has been noted that learning ability is an important variable in experiments of the kind reported here, at least in conditions in which response learning comprises a major part of the task (Warr, 1964*a*, *b*).

It may be argued that differential item difficulty constitutes an artifact in the procedure employed in this series of experiments. It could be the case that all learning is in reality incremental, but that $p(C)$ does not increase regularly with increasing numbers of presentations in these conditions because items presented on several occasions are more difficult than those which are correctly recalled after fewer presentations. A similar argument has been levelled against Rock's (1957) design; the relevant evidence has been discussed by Postman (1963).

It seems unlikely that differential difficulty of items can alone account for the results of experiments, such as Expts. I and II, in which $p(C_{2+})$ does not exceed $p(C_1)$. If this were the case, the same differences in difficulty should have prevented $p(C_{2+})$ from exceeding $p(C_1)$ in the original experiment. Whereas differential difficulty might be expected in general to depress the value of $p(C_{2+})$, this cannot be the whole explanation of the finding that $p(C_{2+})$ is not greater than $p(C_1)$ in certain experiments, since it was greater than $p(C_1)$ in the original experiment when the same items were employed.

There may be more force in the argument from differential difficulty when levelled against experiments in which $p(C_{2+})$ is found to be greater than $p(C_1)$ and in which $p(C)$ does not increase beyond $p(C_2)$. It could be contended that learning here is in reality wholly incremental, but that differential difficulty of the items has prevented an increase being observed beyond $p(C_2)$. If it were the case that this factor were substantially influencing the results of these experiments, however, $p(C)$ ought to be a negatively accelerated function of number of presentations. In other words, the $p(C)$ curve should level off relatively smoothly. An explanation of the results in terms of a difficulty factor of this kind seems unable to account for the marked discontinuity between $p(C_1)$ and $p(C_{2+})$. A more reasonable notion is that an additional factor is exerting a constant influence on $p(C)$ values after the first.

It is therefore concluded on the basis of this series of experiments that task difficulty is an important variable mediating the effect of repetition on learning. In relatively easy conditions learning may occur on an all-or-none basis, but when the level of task difficulty is raised an incremental factor becomes operative. It appears that 'learning' is neither incremental nor all-or-none; rather is learning incremental under some conditions and under others all-or-none.

The author is very grateful to Prof. H. Kay for his comments and assistance in the preparation of this paper.

References

Achilles, E. M. (1920). Experimental studies in recall and recognition. *Arch. Psychol.* **6** (no. 44), 1–60.

Battig, W. F. (1962). Paired associate learning under simultaneous repetition and non-repetition conditions. *J. exp. Psychol.* **64**, 87–93.

Brackett, H. R. & Battig, W. F. (1963). Method of pretraining and knowledge of results in paired associate learning under conditions of repetition and non-repetition. *Amer. J. Psychol.* **76**, 66–73.

CLARK, L. L., LANSFORD, T. G. & DALLENBACH, K. M. (1960). Repetition and associative learning. *Amer. J. Psychol.* **73**, 22–40.

GREGG, L. W., CHENZOFF, A. P. & LAUGHERY, K. R. (1963). The effect of rate of presentation, substitution and mode of response in paired associate learning. *Amer. J. Psychol.* **76**, 110–15.

KAY, H. (1965). One trial learning; incremental or all-or-none? *Acta Psychol.* (In the press).

KRISTOFFERSON, A. B. (1961). Repetition and paired associate learning. *Science*, **134**, 2036–7.

LOCKHEAD, G. R. (1961). A re-evaluation of evidence of one trial associative learning. *Amer. J. Psychol.* **74**, 590–5.

POSTMAN, L. (1962). Repetition and paired associate learning. *Amer. J. Psychol.* **75**, 372–89.

POSTMAN, L. (1963). One-trial learning. Chapter 8 in *Verbal Behaviour and Learning*, ed. Cofer, C. N. and Musgrave, B. S. New York: McGraw-Hill.

POSTMAN, L., JENKINS, W. O. & POSTMAN, D. L. (1948). Experimental comparison of active recall and recognition. *Amer. J. Psychol.* **61**, 511–19.

REED, J. C. & RIACH, W. D. (1960). The role of repetition and set in paired associate learning. *Amer. J. Psychol.* **73**, 608–11.

ROCK, I. (1957). The role of repetition in associative learning. *Amer. J. Psychol.* **70**, 186–93.

ROCK, I. & HEIMER, W. (1959). Further evidence of one trial associate learning. *Amer. J. Psychol.* **71**, 1–16.

UNDERWOOD, B. J., REHULA, R. & KEPPEL, G. (1962). Item selection in paired associate learning. *Amer. J. Psychol.* **75**, 353–71.

WARR, P. B. (1963). The effect of repeated presentation of items on paired associate learning. *Quart. J. exp. Psychol.* **15**, 262–72.

WARR, P. B. (1964*a*). The effect of repetition of items on two phases of paired associate learning. *Quart. J. exp. Psychol.* **16**, 224–31.

WARR, P. B. (1964*b*) Repetition and the recognition of paired associate items. *Quart. J. exp. Psychol.* **16**, 361–3.

(*Manuscript received* 13 *February* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 157–162
Printed in Great Britain

REMINISCENCE, INHIBITION AND CONSOLIDATION

BY S. RACHMAN AND J. GRASSI
Institute of Psychiatry (Maudsley Hospital), University of London

The inability of the two-factor theory of inhibition to account for some of the experimental findings on psycho-motor learning has led Eysenck to formulate a new theory which introduces the concept of consolidation. The present experiment describes an attempt to separate the inhibitory factors and the consolidation process. Fifty-five subjects were randomly assigned to one of four groups and made to practise on the pursuit rotor for five minutes without a rest. They were then re-tested four hours later and reminiscence scores were obtained. The results indicate that if an interfering task is given immediately after the initial practice period is completed, then the process of consolidating the motor learning is adversely affected.

The reminiscence effect, and its relation to massed and spaced practice, have been extensively investigated and most explanations of this phenomenon have been couched in the two-factor theory of inhibition. This theory has proved to be of value and has provided a plausible account of many experimental findings (Kimble, 1949; Eysenck, 1957, 1964; Feldman, 1963). It is now recognized, however, that some experimental results cannot be accommodated by the theory (Adams, 1963; Eysenck, 1964). Here are two examples.

Contrary to prediction, Eysenck and his colleagues found that although high-drive subjects do not perform better than low-drive subjects during the pre-test practice period they nevertheless show greater reminiscence scores (Eysenck & Maxwell, 1961; Eysenck & Willett, 1961). These results are embarrassing because the two-factor explanation of the reminiscence effect assumes the occurrence of pre-rest differences in performance; reminiscence is regarded as an indirect measure of the reactive inhibition which accumulates during pre-rest activity. The dissipation of this I_R during the rest period is said to be responsible for post-rest increments in performance, i.e. the reminiscence effect. A second example is provided by Feldman (1964) who showed that with very long pre-rest practice periods, the reminiscence scores decline. The two-factor theory would of course predict an increase in reminiscence under these conditions.

In an attempt to encompass these and other awkward findings, Eysenck (1964) has proposed a revision and extension of the two-factor theory. The main feature of Eysenck's three-factor theory is the introduction of a concept of consolidation. Reminiscence is defined in the manner of Osgood (1953) as an improvement in performance obtained without further practice. Eysenck argues that this improved performance results from a consolidation of previous learning and from the recovery of performance which was being depressed by the accumulation of inhibition. According to Eysenck, 'different tasks obviously involve different degrees of inhibition and of consolidation in the production of the total effect known as reminiscence'. The pursuit rotor task is offered as an example of a relatively pure 'consolidation' activity, whereas vigilance tasks in which the person's performance is nearly perfect from the beginning are regarded as activities which reflect mainly inhibitory effects.

Eysenck also postulates that 'the memory trace becomes degraded after a certain lapse of time' and that, once degraded, it is no longer available for consolidation. This view appears to imply a gradual decline or degradation in the memory trace; the alternative, a very sudden decline, seems unlikely. One aim of the present experiment was to obtain information about the time course of the postulated memory traces.

Eysenck (1964) uses his three-factor theory not only to account for those findings which cannot be accommodated within the two-factor theory, but also to attempt a re-appraisal of earlier experimental results. This task, is however, complicated by the fact that the early experiments were not designed to separate the processes of consolidation and inhibition dissipation. Consequently, many of the findings can apparently be accounted for by either theory in a satisfactory manner. Rachman (1963), for example, reported that schizophrenic patients only display reminiscence effects on the pursuit rotor after a prolonged rest period. This finding was regarded as evidence of the slow dissipation of inhibition in schizophrenic patients (two-factor theory). It can equally be argued, however, that schizophrenics are very slow to consolidate their learning (three-factor theory). In the circumstances it was felt that an attempt should be made to design an experiment in which the processes of consolidation and I_R dissipation could be teased out.

We accordingly planned an experiment in which the consolidation process is varied while I_R dissipation is held constant. It is known that a 10 min. rest period is sufficiently long to allow for the dissipation of virtually all the I_R which is accumulated during 5 min. of continuous practice on the pursuit rotor, and Ammons (1947) showed that approximately 60% of the I_R dissipates within the first 2 min. of the rest period. In order to ensure that the I_R was fully eliminated in all of our experimental groups we used a 4 hr. rest period. The consolidation process on the other hand was interrupted by giving the subjects a reversed cue task at various intervals after they had completed their 5 min. of pursuit rotor practice.

METHOD

The apparatus and instructions have been given in an earlier paper (Rachman, 1962). The subjects used in this experiment were fifty-five soldiers with a mean age of 21·4 yr (S.D. 2·54). They were randomly allocated to one of four groups. Each subject was given 5 min. of continuous practice on the pursuit rotor, and then re-tested with a further 2 min. of continuous practice after a rest period of 4 hr. The experimental design is shown in Fig. 1.

On completing the 5 min. of practice, the Control Group subjects were given no further tasks or instructions until the re-test 4 hr. later; they were, however, confined to the experimental chamber for the 10 min. immediately following their initial rotor practice. Before starting the re-test each subject was given an abbreviated version of the original instructions.

The subjects of the first Experimental Group (A) were required to practise a reversed-cue task for 3 min. after completing their initial training. They were then confined in the room for a further 7 min. and re-tested on the ordinary pursuit rotor 4 hr. later. In the reversed-cue situation, the subjects had to follow the target while observing the pursuit rotor in a mirror which was placed in an upright position behind the machine. This meant that the target-cue was seen to be rotating in an *anti-clockwise* direction; the subjects were nevertheless required to respond in a clockwise direction. A wooden shield was placed over the rotor to prevent the subjects from getting a direct view of the apparatus.

The subjects in the second Experimental Group (B) rested for 3 min. after the completion of the initial training period, practised the reversed-cue task during the next 3 min. and then rested for 4 min. in the room. They were then re-tested in the usual manner 4 hr. later.

The subjects in the third Experimental Group (C) rested for 6 min. after completion of the initial training, practised the reversed-cue task during the next 3 min. and then rested for 1 min. in the room. They also were re-tested 4 hr. later.

<table>
<tr><th></th><th>5 min.</th><th colspan="4">10 min.</th><th>4 hr.</th><th>2 min.</th></tr>
<tr><td>Control group</td><td>Massed practice</td><td colspan="4">Rest</td><td>Rest</td><td>Massed practice</td></tr>
<tr><td>Exp. group A.</td><td>Massed practice</td><td>Reversed cue (3 min.)</td><td colspan="3">Rest</td><td>Rest</td><td>Massed practice</td></tr>
<tr><td>Exp. group B</td><td>Massed practice</td><td>Rest</td><td>Reversed cue (3 min.)</td><td colspan="2">Rest</td><td>Rest</td><td>Massed practice</td></tr>
<tr><td>Exp. group C</td><td>Massed practice</td><td colspan="2">Rest</td><td>Reversed cue (3 min.)</td><td>Rest</td><td>Rest</td><td>Massed practice</td></tr>
</table>

Fig. 1. The experimental design.

Results

The method of treating the results which was used in earlier work (Rachman, 1962) was repeated in the present experiment. All the practice periods were analysed in terms of 10 sec. trials, thus 5 min. of practice equals 30 trials. Reminiscence scores were obtained in the customary manner by calculating the differences in performance on the last trial, i.e. the last 10 sec. prior to the rest period, and the first trial of the post-rest period. In order to ensure that the reminiscence scores were uninfluenced by group differences in pre-rest performance levels, we carried out an analysis of variance of the pre-rest scores on trial 30 (see Tables 1 and 2). No significant differences

Table 1. *Means and* S.D. *of the pre-rest, post-rest and reminiscence scores for the control and the three experimental groups*

		Pre-rest (trial 30)		Post-rest (trial 31)		Reminiscence	
Group	n	Mean	S.D.	Mean	S.D.	Mean	S.D.
Control	14	1·49	1·299	2·835	1·084	1·34	0·742
A	14	2·21	1·065	2·04	1·027	−0·16	0·837
B	14	1·46	1·069	2·04	1·263	0·59	1·005
C	13	1·52	0·870	1·91	0·963	0·39	0·775

Table 2. *Analysis of variance of pre-rest scores* (*trial* 30)

Source	D.F.	SS	MSV	F	P
Between groups	3	5·39	1·7977	1·5126	N.S.
Within groups	51	60·62	1·1885	—	—
Total	54	66·00	—	—	—

were found. The reminiscence scores on the other hand revealed some significant differences between the groups (see Tables 1 and 3). The mean reminiscence score for the Control Group was 1·34, while Group A which carried out the reversed-cue task immediately had a small but negative score. The remaining Groups B and C obtained small positive scores.

Table 3. *Analysis of variance of reminiscence scores*

Source	D.F.	SS	MSV	F	P
Between groups	3	16·28	5·4267	7·5511	< 0·01
Within groups	51	36·65	0·7186	—	—
Total	54	52·93	—	—	—

The mean score for the Control Group differed significantly from that for Groups A and C ($P < 0{\cdot}01$) and from Group B ($P < 0{\cdot}05$). The means for Groups A and B were also different ($P < 0{\cdot}05$) while that for Group C was not different from either Groups A or B. These results are shown in Fig. 2.

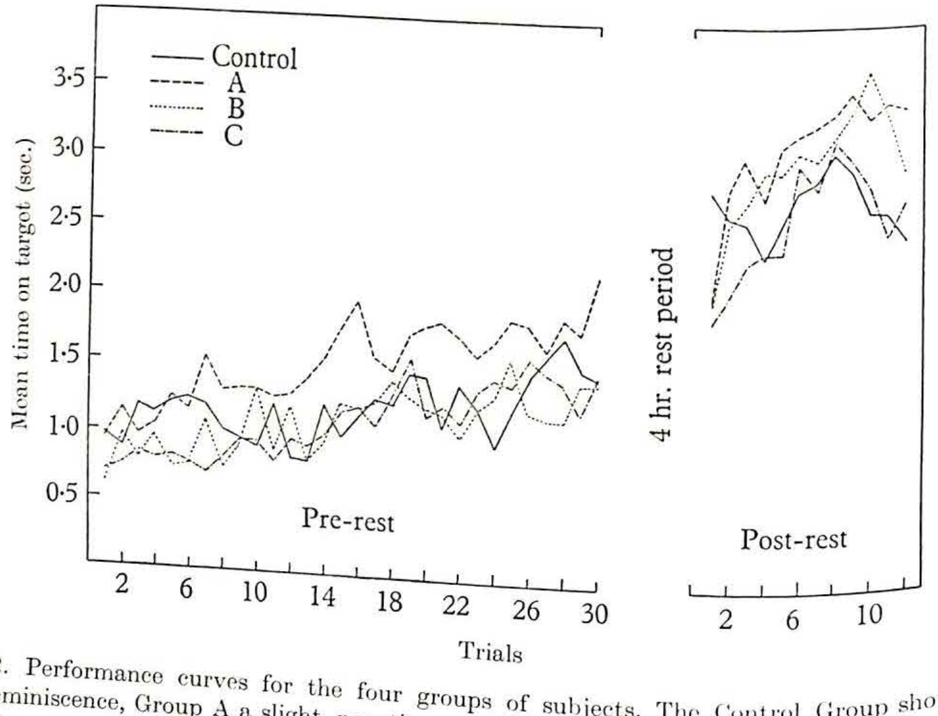


Fig. 2. Performance curves for the four groups of subjects. The Control Group shows normal reminiscence, Group A a slight, negative reminiscence effect and Groups B and C an intermediate amount of reminiscence.

DISCUSSION

The results show that the interposition of a reversed-cue task after pursuit rotor practice brings about a reduction in the reminiscence effect. They also show that the maximum amount of interference occurs when the reversed-cue practice follows immediately after the completion of pursuit rotor practice. This finding lends some indirect support to Eysenck's view that 'the memory trace becomes degraded after a certain lapse of time' in that the consolidation process appears to be particularly susceptible to interference at the end of the 5 min. practice period. Presumably if the reversed-cue practice is given some hours after the completion of the initial training

then the reminiscence score will be unaffected. Our results suggest that the greatest part of the process of consolidating the learning occurs, in this case at least, when the subject stops practising. That the full process of consolidation is not completed within 3 min. is evidenced by the reminiscence scores obtained from the subjects in Groups B and C. The present results provide only an outline of the probable time course of the consolidation process in this learning situation but the experimental design used here can be modified in order to fill in the gaps.

A point of theoretical interest which arises from these results is the demonstration that pursuit rotor reminiscence does not result simply from a recovery from inhibitory effects. In other words, the two-factor theory cannot account for these findings. The subjects, it will be noted, were re-tested after a *four hour* rest period.

It may be argued that the lowered reminiscence scores resulted from the development of a competing response, i.e. moving the stylus in an anti-clockwise direction. This explanation is unlikely for two reasons. First, none of the subjects actually learned how to carry out the mirror task. They tended to flounder about and none of them ever stayed on the target for longer than two-tenths of a second. Secondly, if the poor reminiscence scores were produced by competing responses then we might reasonably expect Groups A, B and C to show similar results as they all had 3 min. of reversed-cue practice. As we have seen, however, Group A gave different results from those of Groups B or C. These results clearly indicate the importance of temporal factors in suppressing the reminiscence scores. As Osgood (1953) has pointed out, 'another implication of a pure interference theory is that amount of forgetting should be *independent of the temporal point of interpolation*' (original italics).

It is, however, necessary to bear in mind that interference and consolidation processes may be expected to interact and that it is rather difficult to separate their respective effects (Bugelski, 1956). The most satisfactory method of separation would appear to be what Bulgelski calls 'suspended animation' (e.g. sleep, freezing, temporary loss of consciousness). The value of this technique is limited however by inherent practical difficulties; some problems cannot be investigated by these methods and need the approach of the present experiment.

References

Adams, J. A. (1963). Comments on Feldman's Reconsideration of the extinction hypothesis of warm-up in motor behaviour. *Psychol. Bull.* **60**, 460–3.

Ammons, R. B. (1947). Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.* **37**, 393–411.

Bugelski, B. R. (1956). *The Psychology of Learning*. New York: Holt.

Eysenck, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge & Kegan Paul.

Eysenck, H. J. (1964). An experimental test of the 'inhibition' and 'consolidation' theories of reminiscence. *Life Sci.*, **3**, 175–88.

Eysenck, H. J. & Maxwell, A. E. (1961). Reminiscence as a function of drive. *Brit. J. Psychol.* **52**, 43–52.

Eysenck, H. J. & Willett, R. A. (1961). The measurement of motivation through the use of objective indices. *J. ment. Sci.* **107**, 961–8.

Feldman, M. P. (1963). A reconsideration of the extinction hypothesis of warm-up in motor behaviour. *Psychol. Bull.* **60**, 452–9.

Feldman, M. P. (1964). Pursuit rotor performance and reminiscence. In *Experiments in Motivation*, Eysenck, H. J. Oxford: Pergamon.

KIMBLE, G. A. (1949). An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.* **39**, 15–23.

OSGOOD, C. E. (1953). *Method and Theory in Experimental Psychology.* New York: Oxford University Press.

RACHMAN, S. (1962). Disinhibition and the reminiscence effect in a motor learning task. *Brit. J. Psychol.* **53**, 149–57.

RACHMAN, S. (1963). Inhibition and disinhibition in schizophrenics. *Arch. gen. Psychiat.* **8**, 91–8.

(*Manuscript received* 24 *April* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 163–181
Printed in Great Britain

A THREE-FACTOR THEORY OF REMINISCENCE

By H. J. EYSENCK
Institute of Psychiatry, University of London

A three-factor theory of reminiscence is suggested, making use of the concepts of consolidation, reactive inhibition and conditioned inhibition. It is further suggested that the reminiscence phenomenon is highly task-specific, in the sense that different tasks call differentially for the various processes hypothesized. Furthermore, it is suggested that differences in drive conditions, personality, fatigue, drug administration and many other variables impose definite limits to the replication of research findings, and that only specific studies of the influence of these variables, within a given theoretical context, can lead to a proper quantitative theory of reminiscence.

Introduction

Reminiscence is usually defined in terms of increments in *learning* which occur during a rest period (Hovland, 1951, p. 653): this author warns that before reminiscence 'can be considered a fundamental *learning* phenomenon, explanation of it in terms of fatigue, motivation, and artifacts of measurement must be eliminated'. Osgood (1953), on the other hand, defines reminiscence as 'a temporary improvement in *performance*, without practice', and: 'The term "reminiscence" refers to the objective fact of improved performance' (p. 509, our italics). It is true that learning is usually indexed in terms of performance, and to that extent the two definitions may be considered equivalent, but it is also true that modern learning theory makes a radical distinction between learning and performance; learning may or may not issue in performance, depending on various conditions which require careful investigation. Some of these conditions are indeed mentioned by Hovland in the sentence quoted above, but the terms used are not precise enough to carry much meaning. Would Hull's concept of 'reactive inhibition' be considered equivalent to 'fatigue', or would it be considered as 'negative motivation'? As long as we have no agreed definition of terms such as these, there might be difficulties in the way of unambiguously demonstrating the phenomenon under investigation. Furthermore, to recognize 'artifacts of measurement' implies knowledge of the true principles of measurement; there is no agreement on just how measurement ought to proceed.

Ammons (1947*a*), to take but one example, has suggested a correction for 'warm-up decrement' which depends for its plausibility on the interpretation of the post-rest improvement in performance *after* the first trial as 'warm-up'; if Eysenck's (1956*b*) explanation of this phenomenon in terms of extinction of conditioned inhibition is preferred, the 'correction' is seen as an artifact which distorts measurement. It can also be shown that results of experiments may be influenced quite powerfully by the definition of single trials used; in the case of the pursuit rotor, for instance, Ammons (1947*b*) has used 1 min trials, Adams & Reynolds (1954) 30 sec trials, and Eysenck (1956*b*) 10 sec trials; in view of the marked changes in performance during the first minute or two of post-rest practice such apparently unimportant differences in choice of trials length may lead to quite different results. As an example, Fig. 1 is taken from a study by Eysenck & Willett (1961); it will be seen that, when the

performance of the high-drive and the low-drive groups is plotted in terms of 10 sec trials, the low-drive group has a reminiscence score which is significantly lower than that of the high-drive group. If the results were plotted in terms of 1 min periods instead, no differences in reminiscence would be apparent. It may be suggested that all results should be reported in terms of 10 sec trials; these short trials could always be statistically combined later, if inspection demonstrated that no information was lost by doing so, whereas, if only relatively long trials are recorded, there is no way in which information on shorter trial lengths could be recovered.

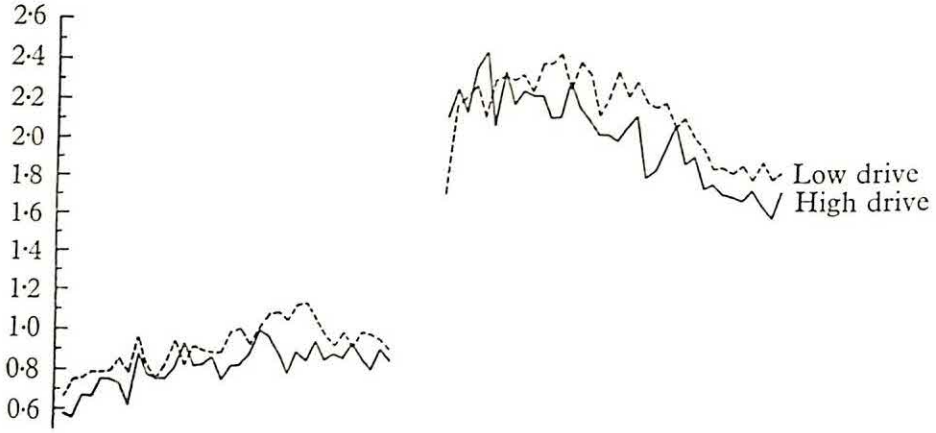


Fig. 1. Reminiscence in high- and low-drive groups. Highly significant differences are obtained when scoring is done in 10 sec periods but these would disappear if scoring were done over 1 min periods (from Eysenck & Willett, 1961).

The difference in the definition of reminiscence between 'increments in learning' and 'increments in practice' becomes important because these differences are to some extent tied up with two sets of theories which have usually been considered antagonistic. It will be suggested that theories of the Hull (1943)–Kimble (1949)–Ammons (1947*a*) type, involving such concepts as reactive inhibition and conditioned inhibition, are closely identified with a definition of the reminiscence phenomenon involving performance and performance decrement, while theories involving such concepts as consolidation and perseveration (Eysenck, 1964*a*) lead rather to definitions involving learning and the neural fixation of learning. It will further be suggested that *both* these sets of theoretical concepts are required to explain the facts of reminiscence, so that instead of being considered *alternative* explanations they should rather be regarded as being *complementary*. Finally, it will be suggested that the degree to which reminiscence is a learning or a performance phenomenon, and is therefore subject to explanation in terms of consolidation or inhibition, depends very much on the task in question; theories of reminiscence are *task-specific*, and it is dangerous to extrapolate hypotheses beyond the particular tests used. We shall in the main be concerned with pursuit-rotor learning, but will occasionally extend our discussion to other types of performance tests; verbal learning and nonsense-syllable learning are excluded from our discussion because of the great difficulties which seem to attend the very demonstration of reminiscence in their field. It is not unlikely that a further principle (interference) plays a much greater part in verbal learning than in the type of behaviour with which we are here concerned.

Inhibition theories of reminiscence

In the past twelve years or so, Hullian theories have been widely used as explanations of reminiscence phenomena, and the associated ones attending massed practice as opposed to spaced practice. The phenomena in question are as follows.

(1) Massed practice results in performance increments which are inferior to those achieved by spaced practice (e.g. Eysenck, 1956*b*).

(2) The superiority of spaced practice is within limits proportional to the length of rest pauses between periods of practice (e.g. Adams, 1954).

(3) Shift from massed to spaced practice, or vice versa, leads to shift in performance such as to make the shifted group resemble the other group more and more, and finally become indistinguishable from it (e.g. Adams & Reynolds, 1954).

(4) Programmed rest periods produce reminiscence (improved performance) in groups with massed practice, but not in groups with sufficiently spaced practice. Even after reminiscence, the performance of groups with massed practice does not usually reach that of groups with spaced practice (e.g. Eysenck, 1956*b*). (Ammons (1947*a*) has introduced the convenient nomenclature of 'temporary work decrement' for that part of the inferiority of the massed practice group to the spaced practice group which is made good during rest; the remainder he calls 'permanent work decrement'. As we have seen under (3) above, the permanent work decrement is not really 'permanent', of course).

(5) Reminiscence is a negatively accelerated function of length of pre-rest massed practice (e.g. Ammons, 1947*b*). (Actually this statement is not quite accurate: reminiscence first increases and then decreases as a function of amount of pre-rest practice; cf. Ammons, 1947*b*; Irion, 1949; Adams & Reynolds, 1954; Feldman, 1964*a*.)

(6) Reminiscence is a negatively accelerated function of length of rest period (e.g. Ammons, 1947*b*).

There are in addition a few phenomena which appear during the post-reminiscence period, and which must be considered in connexion with it; they are dependent on the sequence of rest following massed practice, and do not appear when spaced practice is used, or when massed practice is used without the interposition of a rest period. These phenomena are:

(7) The first post-rest trial (which is used to define reminiscence in conjunction with the last pre-rest trial) is followed by a rapid upswing in performance; this is sometimes referred to as 'warm-up', but will here be called 'post-rest upswing' in order to avoid the theoretical implications of the former term (e.g. Ammons, 1947*a*).

(8) Post-rest upswing is followed by post-rest downswing, i.e. a steady trend of performance in the downward direction. Ammons (1947*a*) suggests that this downward trend ends, and is reversed, when it reaches the point which a comparable massed practice group would have reached if it had not been given a rest period. Denny's (1951) results give some support to this contention, but Ammons's own data (1947*b*) do not. He used thirty-five groups of subjects in all, combining seven lengths of rest period ($\frac{1}{3}$, 2, 5, 10, 20, 60 and 360 min) and five lengths of pre-rest practice periods

($\frac{1}{3}$, 1, 3, 8, and 17 min), and claims that 'recovery after...decline can be easily identified in 16 of the 35 curves...and presumably would have occurred in all if post-rest practice had continued beyond 8 min'. From Ammons's data it is possible to plot the post-rest downswing (and of course the preceding upswing) of the groups subdivided into the seven different 'rest' groups and the five different 'practice' groups. Regardless of what might have happened if practice had continued, the fact remains that, if sixteen of the thirty-five groups show recovery after decline, nineteen show continued decline; this does not suggest that Ammons is justified in claiming that his results support his thesis. The data suggest rather the following generalizations: (8*a*) Groups with short rest periods ($\frac{1}{3}$, 2 min) show little post-rest upswing or post-rest downswing. (8*b*) Groups with longer rest periods show marked post-rest upswing and post-rest downswing, but no trace of a reversal. (8*c*) Groups with short pre-rest practice ($\frac{1}{3}$, 1 min) show neither post-rest upswing nor post-rest downswing. (8*d*) Groups with longer pre-rest practice show both post-rest upswing and post-rest downswing, but no trace of reversal.

Kimble (1949) uses Hull's concept of *reactive inhibition* (I_R) to account for the temporary work decrement, while Hull's concept of *conditioned inhibition* (${}_SI_R$) is used to account for permanent work decrement. According to his account, massed practice produces I_R, which, being a negative drive, impedes performance; hence the inferiority of massed to spaced practice. I_R dissipates during rest, hence the phenomenon of reminiscence. When I_R has grown to be equal to D, the positive drive under which the subject is working during pre-rest massed practice, performance stops and we have a 'block' (Bills, 1931, 1964) or an involuntary rest pause (I.R.P.; Eysenck, 1957); during this I.R.P. I_R dissipates and thus allows performance to begin again. I.R.P.'s are reinforcing, and, as they occur when the subject is resting from the task in question, he is being conditioned not to work in the total stimulus situation of the particular task; hence the habit of not working (${}_SI_R$) is being established. Habits do not dissipate during rest; hence ${}_SI_R$ gives rise to permanent work decrement. However, habit can be extinguished, and Denny, Frisbey & Weaver (1955) have argued that the shift from massed to spaced practice, by eliminating massing (the unconditioned stimulus), should lead to the extinction of ${}_SI_R$ (the conditioned response). Along similar lines Eysenck (1956*b*) has tried to explain the post-rest upswing as being due to extinction of ${}_SI_R$; according to him the UCS is the massing condition of practice including the I.R.P.'s—when these I.R.P.'s are missing in the immediate post-rest period, owing to the dissipation of I_R, extinction must occur. Thus this set of hypotheses would explain all the phenomena in question, with the exception of the post-rest downswing; this remains quite mysterious on any theory hitherto proposed.

This general theory, and the various parts thereof, have been criticized by Koch (1954), Gleitman, Machmias & Neisser (1954), Adams (1961), Jensen (1961) and others; it has been defended by Feldman (1963), among others, who specifically answered certain arguments put forward by Adams (1961). There would appear to be little point in entering into this controversy at this point, but two issues require to be mentioned.

(*a*) In using the Hullian formulation, Eysenck (1957) has made one important change in his conceptualization which appears to be dictated by the pressure of

experimental investigations. Hull accepts the Mowrer–Miller 'work hypothesis', according to which inhibition is a function of the actual physical work done by the organism. Bilodeau (1952), Ellis, Montgomery & Underwood (1952), Bilodeau & Bilodeau (1954), and others, have adduced convincing evidence to the contrary. The writer prefers, as does Walker (1958), a central rather than a peripheral type of hypothesis, relating inhibition to the amount of continued attention required by the task (i.e. a 'mental work' hypothesis rather than a physical one). Evidence for the existence of reminiscence effects in almost purely perceptual tasks (C. H. Ammons, 1955), and studies of bilateral transfer effects (Ammons & Ammons, 1951; Grice & Reynolds, 1952) further serve to discredit the peripheral hypothesis.

(*b*) It is sometimes said that there is no direct evidence for concepts such as I_R and ${}_SI_R$, and Gleitman *et al.* (1954) have, for instance, deduced certain experimental consequences from Hull's postulate of conditioned inhibition which they seem to regard as so unlikely that in the absence of experimental enquiry they are prepared to throw overboard the theory. In putting this hypothesis to the experimental proof, Kendrick (1960) was able to show that the predicted consequences did in fact occur, thus furnishing us with positive proof for the existence of a mechanism very closely resembling ${}_SI_R$ in its operation. Similarly, I.R.P.'s have been identified by Bills (1931) and more recently Spielmann (1963) and Eysenck (1964*d*). This identification is important as the concept of conditioned inhibition stands or falls with the presence of I.R.P.'s in massed practice.

It would of course have been preferable if I.R.P.'s had been identified in pursuit-rotor work, rather than in different activities; Ammons, Ammons & Morgan (1958) have failed to find any such direct support, as have several workers in our own laboratories. The reason would appear to be the high level of time off target on the pursuit rotor, which is confounded with any I.R.P.'s which may occur. Statistical analysis can only disentangle these two sources of error if I.R.P.'s occurred in some regular or rhythmic fashion. The Spielman and Eysenck studies suggest that this is not so, and that I.R.P.'s occur in a random fashion. This difficulty would therefore appear to be almost insuperable in any learning task, such as the pursuit rotor, where an identical index is used for I.R.P.'s and failure to perform perfectly.

We may note at this point also that consolidation phenomena have been independently verified (Glickman, 1961), so that this concept also is not introduced *ad hoc* to serve the purpose of giving the semblance of a proper theoretical interpretation. The position is exactly the opposite: what is known of the workings of the human brain demands that I.R.P.'s, ${}_SI_R$ and consolidation should occur in massed practice on the pursuit rotor, and our task is to use these theoretical concepts to their best advantage.

The partial failure of the inhibition theories

In addition to being able to account for the phenomena discussed in the preceding section, inhibition has been used to make two additional predictions, relating to motivation and to personality. Kimble (1950) suggested that subjects working under conditions of high motivation should show greater reminiscence than subjects working under conditions of low motivation, and his work and that of Wasserman (1951),

Eysenck & Maxwell (1961), Eysenck & Willett (1961), Willett & Eysenck (1962) and Feldman (1964*a*) has indeed shown that this prediction is in accordance with the facts. Kimble's hypothesis was predicated on the assumption that I_R was, as Hull had postulated, a negative drive state; subjects working under a high drive (D) would be able to tolerate a high degree of I_R, and would thus be able to dissipate more I_R during rest.

Eysenck (1956*a*) suggested that extraverted subjects should show greater reminiscence than introverted subjects, and some twenty studies have since been carried out to investigate this postulated relationship between personality and reminiscence (Eysenck, 1962*b*). The great majority have given positive results, although the degree of relationship found tended on the whole to be rather low. Eysenck derived his prediction from the general hypothesis that extraverts would be characterized by greater inhibitory cortical potentials, introverts by greater excitatory cortical potentials (Eysenck, 1957). An excellent discussion of the neurophysiology of inhibition is available in Diamond, Balvin & Diamond (1963), and some recent direct evidence on the relation between personality and inhibition is given by Savage (1964) with respect to E.E.G. patterns, Shagass & Schwartz (1963) with respect to evoked potentials, and Claridge & Herrington (1960) with respect to sedation thresholds; Eysenck (1963) has recently related his conception of inhibition and excitation to the activity of the ascending reticular formation.

While at first the confirmation of these two hypotheses might appear to strengthen the inhibitory theory of reminiscence, it can be shown that the details of the experiments in question do not in fact support the theory. Taking the motivation experiments first, it can be stated that the theory demands that the pre-rest performance of the high-drive group should be superior to that of the low-drive group, at least initially; after pulling even further ahead from the moment that I_R in the low-drive group begins to produce I.R.P.'s, the two curves should run parallel from the moment that I_R in the high-drive group also begins to produce I.R.P.'s. Thus at the end of the pre-rest period there should be a clear-cut superior performance on the part of the high-drive group; the greater degree of I_R tolerated by this group, because balanced by the greater drive, should then allow it to dissipate more I_R during rest, thus producing greater reminiscence. The extensive studies by Eysenck & Maxwell (1961), Eysenck & Willett (1961), Willett & Eysenck (1962) and Feldman (1964*a*) have shown that as far as pre-rest performance is concerned there are no differences between high-drive and low-drive groups; the substantial differences in reminiscence found were all due to post-rest differences in performance. This finding is incompatible with an inhibition hypothesis.

In the case of the relation between personality and reminiscence, a similar condition obtains. According to the hypothesis, greater inhibition in the extraverted group should produce a greater performance decrement; the dissipation of this greater performance decrement would then show up in the form of greater reminiscence. Most studies have simply correlated reminiscence with extraversion, but, while the positive coefficients usually found would certainly be compatible with the hypothesis, they might with equal ease have resulted from a set of scores where pre-rest performance was equal, but post-rest performance favoured the extraverts. A special study to investigate this point was carried out by Eysenck (1964*b*), who showed that

this second possibility was much closer to the facts; he failed to find any pre-rest difference in performance between his extraverted and introverted subjects. This finding too is incompatible with the inhibition hypothesis.

A third prediction was made from the inhibition hypothesis by Rachman (1962), and here too the prediction was confirmed, but the details of the confirmation served to discredit the inhibition hypothesis. Rachman argued that any strong 'alien' stimulus, such as a loud buzzer, if applied shortly before the rest period on a massed-practice pursuit-rotor task, should have the effect of *disinhibiting* part of the I_R accumulated up to this point; this would improve performance and lower reminiscence. This lowering of reminiscence was indeed found, both by himself and by Feldman (1964*b*), but the effect of the alien stimulus was not to raise performance pre-rest, but rather to lower it post-rest! This result, too, must therefore be counted as disconfirming the inhibition hypothesis. An even more crucial experiment might be one in which the alien stimulus was applied during the rest period, rather than during the pre-rest period; an experiment somewhat along these lines will be discussed in a later section.

These three experiments were performed in order to test predictions made from the basis of the inhibition theory, and while they verified the prediction they did so in a manner which in fact discredited the theory. The next experiment to be discussed was performed with the express intention of testing the inhibition theory directly. Inhibition theory postulates *depression of pre-rest performance* as the crucial factor in reminiscence; reminiscence is due to recovery from this depression. In his experiment, Eysenck (1964*a*) divided 300 subjects into groups equated for initial ability on the pursuit rotor, but either showing or not showing depression of performance during the last 90 sec of pre-rest practice. On the inhibition theory it would be expected that those subjects showing most depression of performance pre-rest (i.e. a depression theoretically due to inhibition) would dissipate most I_R during rest, thus showing greater reminiscence than subjects not showing any pre-rest performance depression. Nothing of the kind was in fact found; reminiscence scores were completely independent of amount of pre-rest performance decrement. While this experiment too is not crucial, it does argue strongly against the inhibition theory.

Consolidation theories of reminiscence

A consolidation or 'perseveration' theory of memory was first put forward by Müller & Pilzecker (1900), although not in relation to reminiscence. According to this theory, a neural fixation process is assumed to continue after the organism is no longer confronted with the set of stimuli which constitute the learning task. This fixation process plays a crucial part in efficient retention, according to the consolidation hypothesis, and anything that interferes with perseveration is assumed to have an adverse effect on the subject's ability to transfer material acquired to the permanent memory store. Between the wars this theory fell into disrepute, and McGeoch & Irion (1952) dismissed it as lacking 'any great generality' in their discussion of theories of reminiscence. However, recent work has rendered it respectable again, and there is now some direct evidence to demonstrate its relevance to reminiscence phenomena.

Work on retrograde amnesia, for instance, is difficult to explain on any other lines (Russell & Nathan, 1946), although, being only clinical, it is of course not well enough controlled to be convincing by itself (Coons & Miller, 1960). Experimental work with electro-convulsive shock is more convincing; it has been shown, both with humans and with animals, that, if shock is given between learning and remembering, then genuine retrograde amnesia is produced. It has further been shown that the longer shock was delayed after learning the less was the resulting amnesia (Flescher, 1941; Zubin & Barrera, 1941; Williams, 1950; Duncan, 1949; Cronholm & Molander, 1958; for review see Campbell, 1960). Anoxia (Hayes, 1953; Thompson & Pryer, 1956) and anaesthesia (Leukel, 1957; but see Russell & Hunter, 1937) are other experimental procedures which have given positive results in this connexion, as has the direct stimulation of certain midbrain structures (Glickman, 1958; Thompson, 1958). The evidence for some sort of consolidation process is thus rather convincing, and there is now even some physiological evidence to suggest in more precise terms the how and where of consolidation (Stellar, 1957; Burns, 1958). Glickman (1961) has furnished a fairly recent review of the evidence, and there are in addition some even more recent studies suggesting the detailed working of the determination of reminiscence by consolidation.

Of particular interest in this connexion is the work of Walker (1958) on what he calls *action decrement*. While much of his own work has dealt with verbal learning and animal studies, general relevance is claimed for the main generalizations arrived at (Walker & Tarte, 1963). These two writers summarize the propositions of their theory as follows: '(1) The occurrence of any psychological event...sets up an active, perseverative trace process which persists for a considerable period of time. (2) The perseverative process has two important dynamic characteristics: (*a*) permanent memory is laid down during this active phase in a gradual fashion; (*b*) during the active period, there is a degree of temporary inhibition of recall, i.e. action decrement (this negative bias against repetition serves to protect the consolidating trace against disruption). (3) High arousal during the associative process will result in a more intensely active trace process. The more intense activity will result in greater ultimate memory but greater temporary inhibition against recall.' We shall attempt to adapt these notions, in a somewhat modified form, to the problems of pursuit-rotor learning; for the moment let us only notice that no form of consolidation hypothesis by itself can suffice to explain the phenomena associated with reminiscence we have listed. Consolidation theory is adequate to explain reminiscence itself, and the differences between massed and spaced learning; it cannot explain permanent work decrement, post-rest upswing, or the facts of shifting from massed to spaced practice.

When a theory is clearly not self-sufficient to explain a set of phenomena, one is naturally somewhat reluctant to make use of the theory at all; it seems preferable to try and make do with a smaller set of explanatory variables if possible. There is, however, one experimental study which seems to indicate without any doubt the prime importance of consolidation for reminiscence. In this study Rachman & Grassi (1965) used one control group and three experimental groups. All groups practised for 5 min on the pursuit rotor under conditions of massing, and all groups were given a rest period of 4 hr during which they left the laboratory, before being retested for reminiscence. Equally, all groups were retained for 10 min immediately following the

pre-rest practice. During this period, the control group rested, while the three experimental groups practised on a reversed-cue (mirror) pursuit rotor, it being hypothesized that this practice would interfere with consolidation. One of the experimental groups (group A) carried out this practice during the first 3 min after the 5 min pre-rest practice; the next experimental group (group B) carried out this practice during the 4th to the 6th min, while the third experimental group (group C) carried out this practice during the 7th to the 9th min. Inhibition theory would predict that all four groups would have identical reminiscence scores, 4 hr of rest being quite adequate for all inhibition acquired during the 5 min of practice to dissipate. Consolidation theory, on the other hand, would predict most reminiscence for the control group, least for group A, with groups B and C intermediate. The results bore out the prediction derived from the consolidation theory at an acceptable level of statistical significance.

Of equal interest is a set of experiments which, although not using pursuit-rotor practice or even human subjects, seems to have given strong support to the consolidation theory of memory. In a series of studies McGaugh & Petrinovich (1959), McGaugh, Westbrook & Thomson (1962), and Breen & McGaugh (1961) have injected stimulant drugs into rats after completion of learning periods, and tested the rats after the drug effects had worn off; comparison with control groups demonstrated the superiority of the drug-treated groups, and the authors concluded that the experiments could best be interpreted as showing that drug administration 'improves maze performance by facilitating post-trial consolidation of the neurophysiological process underlying memory' (McGaugh *et al.* 1962, p. 172). Some criticisms have been made of this work (Thiessen, Schlesinger & Calhoun, 1961), but McGaugh & Petrinovich (1963) have succeeded in answering these satisfactorily; it would appear therefore that we must accept this evidence in favour of the consolidation hypothesis.

A combined inhibition–consolidation theory of reminiscence

The theory to be suggested here combines the essential features from the Kimble two-factor (inhibition) theory and the consolidation hypothesis. We may reconstruct the course of events during pursuit-rotor learning somewhat as follows. (i) During pre-rest practice, I_R builds up and finally enforces I.R.P.'s; the point at which I.R.P.'s begin to occur depends on the drive level under which the subject is working. No permanent memory traces are laid down, and hence no learning takes place. (This statement may require qualification; I.R.P.'s may provide occasions for laying down permanent memory traces, but the very short periods in question are not likely to influence our argument to any great extent.) I.R.P.'s provide the reinforcement for the growth of ${}_SI_R$, but this also, being a habit, fails to lay down permanent memory traces. (ii) A programmed rest pause allows consolidation of the pursuit-rotor habit to take place, following a negatively accelerated curve of acquisition; this provides the basis of the reminiscence phenomenon. The rest pause also allows ${}_SI_R$ to consolidate; this habit too follows a negatively accelerated curve of acquisition. The consolidation of ${}_SI_R$ provides the basis for the permanent work decrement. (iii) Resumption of work after the rest pause produces extinction of ${}_SI_R$, due to non-reinforcement; ${}_SI_R$ begins to accumulate again once sufficient I_R has been built up

to produce I.R.P.'s. Working against the post-rest upswing produced by this extinction process is the still-continuing consolidation process; as our quotation from Walker and Tarte has made clear, we conceive of consolidation and work as mutually interfering processes. This interference produces post-rest downswing, which in turn ceases when consolidation is complete; at this point we may then return to the gentle upward-sloping course characteristic of massed practice without rest pause interference.

This theory would seem to account for all the phenomena listed above (p. 165). How does it handle the phenomena of personality and motivation? With respect to the former phenomenon, Eysenck (1964*b*) has suggested that the differences in conditionability characteristic of extraverted and introverted subjects (Eysenck, 1962*a*) may be responsible. Introverts condition quickly and strongly, and accordingly will form more $_sI_R$ in the course of the pre-rest practice period; hence after rest they will show greater permanent work decrement. The extinction of $_sI_R$ after resumption of work will then cause the differences in performance between extraverts and introverts to disappear. The plot of the detailed results given by Eysenck (1964*b*) bears this analysis out in all important particulars. An alternative and rather implausible hypothesis might be that extraverts consolidate better, or learn better in the first place; this hypothesis would be purely *ad hoc* and would in any case not account for the gradual disappearance of the observed differences after resumption of post-rest practice.

As regards motivation, it is to be noted that the major reminiscence effects are lasting, rather than transitory; in other words, the hypothesis that high drive leads to better learning and/or consolidation than low drive is not untenable. No detailed working out of this suggestion will be given here as the point is not central to our argument, and as there are a number of other problems which are in more urgent need of discussion. Normally one would probably have preferred to assume that drive facilitates *learning* rather than consolidation, but the recent work of McGaugh and his colleagues already referred to makes any confident assertion unwise.

Task-specific features of reminiscence

Reminiscence in pursuit-rotor learning is almost entirely due to consolidation; in a task such as tapping (Grassi, 1964) it is almost entirely due to reactive inhibition (Fig. 2)—there is in such a task no learning that could consolidate! Why is it that in one task reactive inhibition has no effect in depressing performance, i.e. on the pursuit rotor, while on another one, such as tapping, there is a very profound effect? Why is it that some tasks, such as the pursuit rotor, are almost unaffected by differences in drive, while others, such as conditioning (Willett, 1964), are very much affected? It would appear that there are several dimensions along which tasks can be ranged, and it may be suggested that an investigation of these dimensions could be of very great importance in understanding the phenomena associated with learning and reminiscence. We may postulate three main dimensions.

(1) In the first place, we have tasks which require new learning, such as pursuit-rotor performance; here consolidation of this new learning is obviously of prime importance. At the other end of the continuum are well-practised tasks not involving

new learning, such as tapping (Grassi, 1964), vigilance (Buckner & McGrath, 1963), and visual after-effects (Holland, 1963). Here there is no consolidation, but only reactive inhibition. Other tasks are intermediate, such as inverted alphabet printing, or the pathways test.

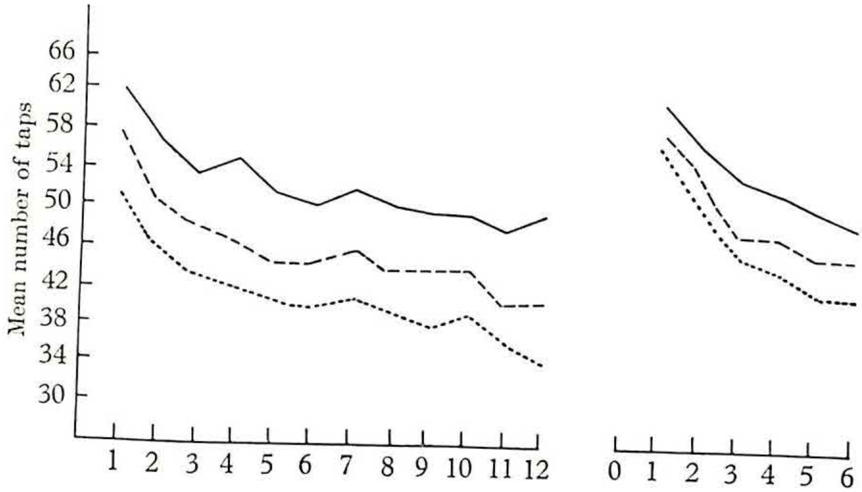


Fig. 2. Reminiscence on tapping task in three groups of subjects (from Grassi, 1964). —, Normal; - - -, behaviour disorder; ..., brain damage.

(2) In the second place, we may order our tasks along a continuum according to the degree to which they are likely to be influenced by the blocks which we may regard as the only demonstrable evidence of reactive inhibition. Self-paced tasks, such as pursuit-rotor performance, would be at one extreme, being almost immune to the effects of inhibition, while experimenter-paced tasks, such as short-time vigilance tests or experimenter-paced reaction time tests, would be at the other. It may be surmised that pursuit-rotor work would be relatively independent of the blocks that might occur, because performance decrements occurring during the block could be made good by improved performance immediately after the block. It should not be impossible to test this assumption experimentally. Broadbent (1953) has also commented on the fact that blocks sometimes do and sometimes do not produce work decrement, and he has suggested a possible task parameter to account for these differences. In the vigilance tasks studied by him, the signal to be detected may be presented for a long or a short time; if a block occurs during the presentation of a short-time signal, this will be missed, and the block will produce a performance decrement. If the block occurs during the presentation of a long-time signal, the signal will still be there after the block has disappeared, and will therefore be noted; there will be no performance decrement. In the case of tasks such as tapping, we may suggest a somewhat analogous mechanism: the block produces a marked slowing down during one of the taps, but this may be made good by a particularly quick series of taps immediately following the block, made possible by the shedding of inhibition which accompanies the involuntary rest pause or block. Thus we have found that extraverts have many more blocks in tapping than do introverts, yet their actual output is equal to that of the introverts (Spielmann, 1963; Eysenck, 1964*d*).

(3) In the third place, we must turn to the effects of drive. Tasks differ from each other along a continuum, the one end of which is characterized by experimental operations which are of an all-or-none character—you either carry them out properly, or not at all. An intelligence test is perhaps a good example—provided an individual agrees to carry out the test at all, degree of motivation does not seem to have any great influence on his performance (Eysenck, 1944; Tiber & Kennedy, 1964). Other tasks are infinitely variable, in the sense that all types of intermediate performance are possible; the serial reaction time test would be a good example of this other extreme of the continuum, and so would tapping. The hypothetical regression lines of 'Type 1' tasks and 'Type 2' tasks are plotted in diagrammatic form in Fig. 3;

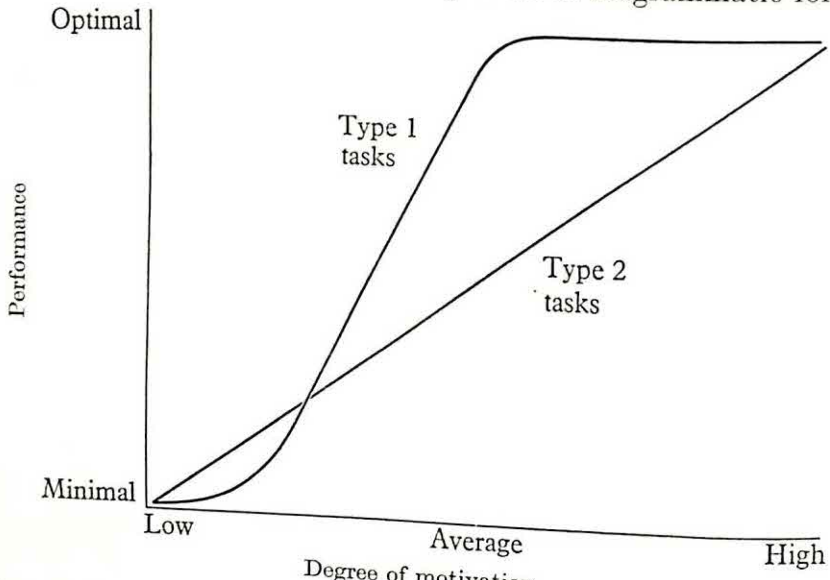


Fig. 3. The relation between performance and degree of motivation in two types of tasks.

they indicate the relative independence of the former, and the close dependence of the latter, on the existing state of drive of the experimental subject. The failure of increased drive to produce greater performance on the pursuit-rotor task would in terms of this analysis be explained as being due to this task being of Type 1; the subject either carries out the task as best he can, or he fails completely to carry it out. The task is not infinitely variable; even a slight general decrement means that the target is hardly ever reached. Type 1 and Type 2 tasks only serve to define the extremes of a continuum, of course; most tasks will be intermediate between these two extremes. The differentiation here intended is very similar to that made by economists when they talk about demand being 'elastic' or 'inelastic' when plotted as a function of price. It should be borne in mind, of course, that other principles, such as the Yerkes–Dodson law, must also be taken into account before making any detailed predictions.

While we may perhaps speak of 'task specificity' in connexion with the role of motivation, or the performance decrements produced or not produced by I_R, it should be remembered that generalizations can only be made for specific populations tested under specified conditions. Brain-damaged subjects, or chronic schizophrenics, may show such low degrees of motivation as might not be found in any

normal group, and for them the relative invulnerability of pursuit-rotor performance to drive might cease to hold true. Conditions of fatigue or drug injection may have effects on normal subjects which temporarily shift their performance to parts of the three-dimensional task model which are outside the boundaries of normal groups. There is some evidence that fatigue produces unusually long I.R.P.'s (Bjerner, 1949; Williams, Lubin & Goodnow, 1959), and, while the usual I.R.P.'s on the pursuit rotor may be too short to depress performance, very long I.R.P.'s may be impossible to correct by greater effort immediately succeeding the rest pauses. All these qualifications must be held in mind when making predictions in relation to any specific task for any specific population.

It will be clear from what has been said that it is not possible to speak of a theory of reminiscence in any meaningful fashion; reminiscence is not a single phenomenon with a single explanation, but rather a broad descriptive term covering several different phenomena. Pursuit rotor reminiscence, in terms of our theory, is due to consolidation; reminiscence in tapping, or vigilance, in terms of our theory, is due to the dissipation of I_R; reminiscence in rotating spiral after-effects, in terms of our theory, is due to the dissipation of ${}_SI$ (stimulus satiation). Other tasks may combine these different mechanisms in varying proportions, or, as in the case of verbal learning, introduce other mechanisms, such as interference. Generalization across tasks is clearly dangerous and difficult.

A QUANTITATIVE MODEL OF PURSUIT-ROTOR REMINISCENCE

Several attempts have been made, notably by Kimble & Shatel (1952), to quantify such concepts as I_R and ${}_SI_R$ in connexion with pursuit-rotor learning, and demonstrate their growth curves with changes in duration of rest and duration of pre-rest practice; similarly, Eysenck & Willett (1961*a*) and Feldman (1964*a*) have plotted the growth of I_R as a function of drive. All this work was done with the explicit assumption that reminiscence was a function of inhibition; if it is admitted that reminiscence is instead a product of consolidation, then it might be thought that the curves and formulae hitherto used to link reminiscence and I_R may now be used to give a quantitative formulation of the growth and decline of the consolidation process instead. (Work on ${}_SI_R$ is not implicated in this change of theory, as it is still assumed that permanent work decrement is due to conditioned inhibition; provided that consolidation is allowed to proceed in full, the failure of the reminiscence effect to reach the level of a comparable spaced practice group is an adequate measure of ${}_SI_R$.)

However, there are clearly some difficulties in the way of any simple transformation. Consider Fig. 4, which shows the growth of reminiscence as a function of pre-rest work period duration, for groups working under high and low drive respectively. It would seem possible to regard these curves as measures of *learning*; longer pre-rest work periods give rise to greater learning, which approaches asymptotic values dependent on the drive under which the subject is working. However, as pointed out before, there is also the possibility that the consolidation process is affected by drive, so that the shapes of the resulting curves might be the joint effects of (*a*) degree of learning, and (*b*) amount of consolidation. Until this problem is resolved we cannot

make any straightforward identification between theoretical variables and observable values.

There is one observation which points to the possible relevance of consolidation in determining the shape of the curve of reminiscence as a function of pre-rest practice. Ammons (1947*b*), Feldman (1964*a*), and others, have found that this curve, once it has reached a plateau, tends to decline again; thus reminiscence after 20 min of practice is less than after 5 min. It is possible to account for this curious fact along the following lines. (*a*) Learning follows the course of a negatively accelerated exponential function; i.e. most learning takes place in the first few minutes. (*b*) Consolidation, once it is started during a rest pause, continues to run its course, until it ceases after a definite period of time has elapsed. (*c*) Memory traces are available for consolidation only for a limited period of time. Let us assume that memory traces are in either one or the other of two states, i.e. either available or not available, and

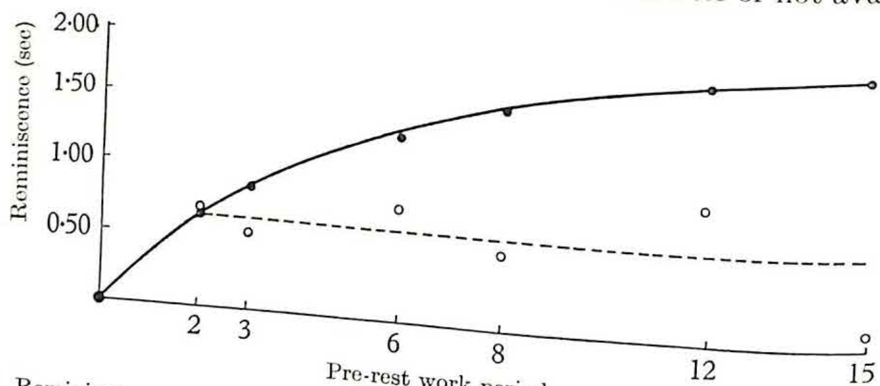


Fig. 4. Reminiscence as a function of pre-rest work period in high- and low-drive groups (from Willett & Eysenck, 1962). ●—●, High drive; ○---○, low drive.

that they cease to be available after 15 min. This means that a rest pause introduced after 15 min of practice is just in time to permit all the accumulated memory traces to consolidate and enter the permanent memory storage. A rest introduced after 16 min will exclude all memory traces laid down during the first minute of practice; but this minute has produced the greatest amount of learning (cf. (*a*) above). Consequently the total amount of learning transferred to permanent memory storage will be less than in the case of the 15 min practice period, and accordingly the curve plotted in Fig. 4 will begin to decline. As pre-rest practice is extended more and more, the decline will continue, because more and more the large amounts of learning that have taken place during the first few minutes will become unavailable, and the relatively small amounts of learning that have taken place during the last few minutes will take their place. This hypothesis may have to be amplified to take into account the possibility that memory traces can exist in more than two states, but within its limitations it does account for a very mysterious effect. If this account be accepted, then we could use the hypothesis in turn to investigate along quantitative lines the disappearance of memory traces under conditions where consolidation is made impossible.

Having in our interpretation lost the possibility of plotting the progress of I_R growth directly, we must search for alternative ways of measuring or indexing this

variable. One possibility is indicated in an early study by Eysenck (1956*b*), who gave eleven 2 min practice sessions divided by 5 min intervals; as shown in Fig. 5 the post-rest upswing and post-rest downswing phenomena are clearly seen in all but the first two and possibly the last trials. On our interpretation the existence of the post-rest upswing demonstrates that during the preceding period enough I_R has developed to produce I.R.P.'s, which in turn cause ${}_sI_R$ to develop; it is the extinction of this ${}_sI_R$ which is shown on the graph as post-rest upswing. (Adams, 1963, has criticised this demonstration, but Feldman, 1963, has rebutted these criticisms.) It would seem to follow from our theory that, if the practice periods were shortened sufficiently, I_R would be prevented from building up to a sufficient level to produce I.R.P.'s, and no ${}_sI_R$ could develop; hence under these conditions we should have no post-rest upswing, but only post-rest downswing. Feldman (1964*a*) has reported such an experiment, subjects practising for fifteen 20 sec periods, separated by 40 sec rest

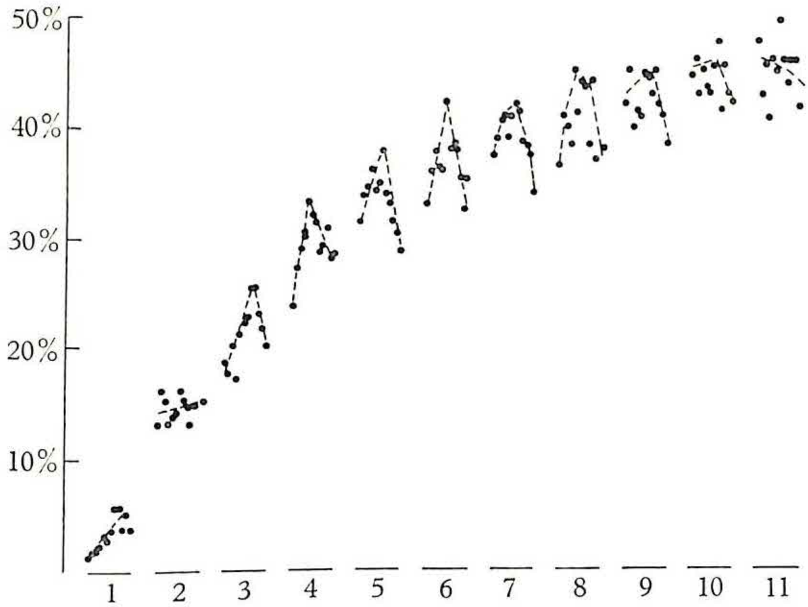


Fig. 5. Post-rest upswing and downswing phenomena. As observed during eleven 2 min practice sessions divided by 5 min intervals (from Eysenck, 1956*b*).

pauses (Fig. 6). It will be seen that our prediction is verified: there is no post-rest upswing on any of the trials, only post-rest downswing. These two practice periods (2 min and 20 sec respectively) thus straddle the moment when I.R.P.'s develop, and repeating the experiment with various intermediate periods should disclose the precise length of pre-rest practice required to produce I.R.P.'s, and consequently post-rest upswing. It may be added, parenthetically, that in Fig. 6 both high-drive and low-drive groups fail to develop post-rest upswing; it follows from our general set of hypotheses that with increase in the length of the pre-rest practice period the low-drive group should show post-rest upswing earlier than the high-drive group. Pre-rest practice of 90 sec would seem to be just on the borderline; Fig. 7 is reproduced from Star (1957), and shows results from sixteen 90 sec work periods separated by 5 min rest periods. It will be seen that post-rest upswing is present but in a very

rudimentary form only. The subjects were students, so that nothing is known about their level of motivation; it would seem that this experiment could with advantage be repeated on groups working under known conditions of high and low motivation.

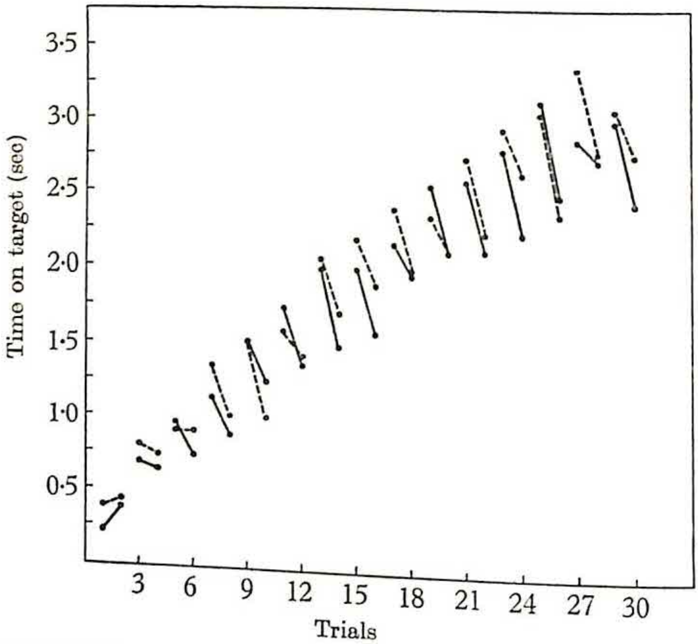


Fig. 6. Post-rest downswing and lack of post-rest upswing as observed during fifteen 20 sec practice periods separated by 40 sec rest periods (Feldman, 1964*a*). - - -, High drive —, low drive.

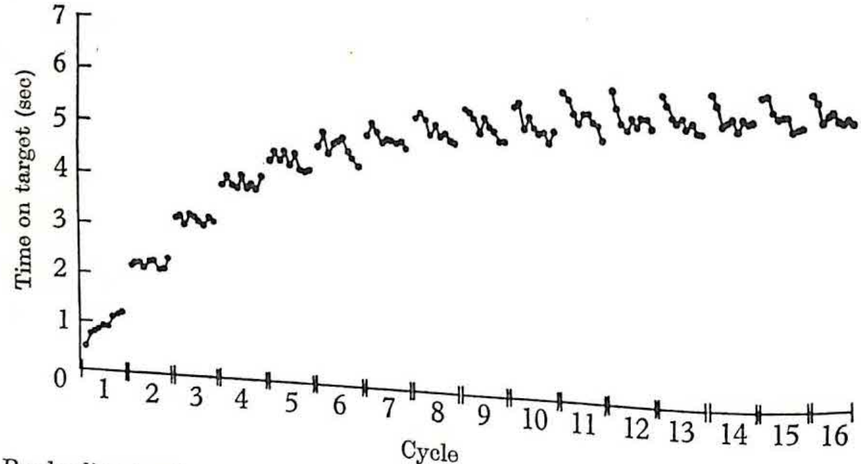


Fig. 7. Border-line development of post-rest upswing and downswing during sixteen 90 sec work periods separated by 5 min rest periods (Star, 1957).

Discussion

The theory tentatively presented here has many weaknesses, which derive from several different sources. Clearly, for instance, what we have to say about the influence of drive on different types of tasks, while receiving some support from such studies as those reported by Eysenck (1964*c*), must remain speculative until the whole theory of motivation is placed on a sounder and more widely acceptable basis.

Similarly, our discussion of personality correlates of reminiscence must remain speculative until there is a greater amount of agreement on the basic and most fundamental concepts and laws in that field. In other words, any theory of reminiscence will be limited by the failure of associated fields in psychology to put their own house in order.

Within the theory itself, the most obvious drawback is the lack of clear quantification of hypothesized variables. Kimble's and Ammons's two-factor theories held out a promise of quantification of I_R and $_sI_R$, but this promise was clearly premature. The greater complication attending greater sophistication in theory makes observable phenomena less likely to serve directly as measures of underlying theoretical concepts. This situation has led some observers, such as Adams (1963), to suggest giving up theory altogether and relying entirely on inductive studies. This would seem to be a counsel of despair; good theories are the end result of a long process of refinement of bad theories progressively improved through checking and testing of deductions, and the belief that good theories will materialize suddenly if only enough inductive work not guided by any theory is carried out seems to lack support in the history of science.

A third difficulty lies in the task-specific nature of many of the concepts involved. If this suggestion of a close relation between type of task and implication of inhibition and/or consolidation be accepted, and if we agree additionally that this relation itself is modified according to the type of person tested, his drive and fatigue state, and any chemical (drug) influence to which he may have been subjected, then it will be clear that any truly quantitative statement of a proper theory of reminiscence is still very much in the future. However, these are difficulties implicit in the subject-matter of psychology, and cannot be shirked or avoided. They may serve to explain the occasional failure of one experimenter to confirm results reported by another; while such variables as size of target, speed of rotation and stance of subject can be controlled (but frequently are not), such variables as the drive or personality of the subject are hardly ever stated by the original experimenter, and would in any case be extremely difficult to measure or control.

In spite of these admitted difficulties, the theory here presented may be useful in suggesting fruitful ways of designing future experiments in this field, of testing the various intertwined strands of the hypotheses involved, and in making clear the enormous complexity of what at first seemed a simple and clear-cut phenomenon.

The writer is indebted to D.S.I.R. for financial support of the research work on which this theoretical generalization is based.

References

Adams, J. A. (1954). Psychomotor performance as a function of intertrial rest interval. *J. exp. Psychol.* **48**, 131–3.

Adams, J. A. (1961). The second facet of forgetting: a review of warm-up decrement. *Psychol. Bull.* **58**, 257–73.

Adams, J. A. (1963). Comment on Feldman's 'reconsideration of the extinction hypothesis of warm-up in motor behavior'. *Psychol. Bull.* **60**, 460–3.

Adams, J. A. & Reynolds, B. (1954). Effect of shift in distribution of practice conditions following interpolated rest. *J. exp. Psychol.* **47**, 32–6.

AMMONS, C. H. (1955). An experimental study of temporary and permanent inhibitory effects associated with acquisition of a simple perceptual–motor skill. Louisville, Kentucky: unpublished Ph.D. thesis.

AMMONS, R. B. (1947*a*). Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.* **54**, 263–81.

AMMONS, R. B. (1947*b*). Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.* **37**, 393–411.

AMMONS, R. B. & AMMONS, C. H. (1951). Bilateral transfer of rotary pursuit skill. *Amer. Psychologist*, **6**, 294 (abstract).

AMMONS, R. B., AMMONS, C. H. & MORGAN, R. L. (1958). Subskills in rotary pursuit as affected by rate and accuracy requirements and by distribution of practice. *J. gen. Psychol.* **58**, 259–79.

BILLS, A. G. (1931). Blocking: a new principle of mental fatigue. *Amer. J. Psychol.* **43**, 230–45.

BILLS, A. G. (1964). A study of blocking and other response variables in psychotic, brain-damaged and personality-disturbed patients. *Beh. Res. Ther.* **2**, 99–106.

BILODEAU, E. A. (1952). Decrements and recovery from decrements in a simple work task with variation in force requirements at different stages of practice. *J. exp. Psychol.* **44**, 96–100.

BILODEAU, I. McD., & BILODEAU, E. A. (1954). Some effects of work loading in a repetitive motor task. *J. exp. Psychol.* **48**, 455–67.

BJERNER, B. (1949). Alpha depression and lowered pulse rate during delayed actions in a serial reaction test: a study in sleep deprivation. *Acta physiol. scand.* **16** (Suppl. 65), 1–93.

BREEN, R. A. & McGAUGH, J. L. (1961). Facilitation of maze learning with post-trial injections of picrotoxin. *J. comp. physiol. Psychol.* **54**, 498–501.

BROADBENT, D. E. (1953). Neglect of the surroundings in relation to fatigue decrements in output. In *Fatigue* (eds. W. F. Floud and A. T. Welford). London: H. K. Lewis.

BUCKNER, D. N. & McGRATH, J. J. (ed.) (1963). *Vigilance: a Symposium*. London: McGraw-Hill.

BURNS, B. D. (1958). *The Mammalian Cerebral Cortex*. London: Arnold.

CAMPBELL, D. (1960). The psychological effects of cerebral electroshock. In *Handbook of Abnormal Psychology* (ed. H. J. Eysenck). London: Pitman.

CLARIDGE, G. S. & HERRINGTON, R. N. (1960). Sedation threshold, personality, and the theory of neurosis. *J. ment. Sci.* **106**, 1568–83.

COONS, E. E. & MILLER, N. E. (1960). Conflict versus consolidation of memory traces to explain 'retrograde amnesia' produced by E.C.S. *J. comp. physiol. Psychol.* **53**, 524–31.

CRONHOLM, B. & MOLANDER, L. (1958). Influence of an interpolated E.C.S. on retention of memory material. *U. Stockholm psychol. Lab. Rep.* no. 61.

DENNY, M. R. (1951). The shape of the post-rest performance curve for the continuous rotary pursuit task. *Mot. Skills Res. Exch.* **3**, 103–5.

DENNY, M. R., FRISBEY, N. & WEAVER, J. (1955). Rotary pursuit performance under alternate conditions of distributed and massed practice. *J. exp. Psychol.* **49**, 48–54.

DIAMOND, S., BALVIN, R. S. & DIAMOND, F. R. (1963). *Inhibition and Choice*. New York: Harper and Row.

DUNCAN, C. P. (1949). The retroactive effect of electroshock on learning. *J. comp. physiol. Psychol.* **42**, 32–44.

ELLIS, D. S., MONTGOMERY, V. & UNDERWOOD, B. J. (1952). Reminiscence in a manipulative task as a function of work surface height, pre-rest practice, and interpolated rest. *J. exp. Psychol.* **44**, 420–7.

EYSENCK, H. J. (1944). The effect of incentives on neurotics, and the variability of neurotics as compared with normals. *Brit. J. med. Psychol.* **20**, 100–3.

EYSENCK, H. J. (1956*a*). Reminiscence, drive and personality theory. *J. abnorm. soc. Psychol.* **53**, 328–33.

EYSENCK, H. J. (1956*b*). 'Warm-up' in pursuit rotor learning as a function of the extinction of conditioned inhibition. *Acta psychol., The Hague*, **12**, 349–70

EYSENCK, H. J. (1957). *Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.

EYSENCK, H. J. (1962*a*). Conditioning and personality. *Brit. J. Psychol.* **53**, 299–305.

EYSENCK, H. J. (1962*b*). Reminiscence, drive and personality—revision and extension of a theory. *Brit. J. soc. clin. Psychol.* **1**, 127–40.

EYSENCK, H. J. (1963). The biological basis of personality. *Nature, Lond.*, **199**, 1031–4.

EYSENCK, H. J. (1964*a*). An experimental test of the 'inhibition' and 'consolidation' theories of reminiscence. *Life Sci.* **3**, 175–88.

EYSENCK, H. J. (1964*b*). Personality and reminiscence—an experimental study of the 'reactive inhibition' and the 'conditioned inhibition' theories. *Life Sci.* **3**, 189–98.

EYSENCK, H. J. (ed.) (1964*c*). *Experiments in Motivation.* Oxford: Pergamon.

EYSENCK, H. J. (1964*d*). Involuntary rest pauses in tapping as a function of drive and personality. *Percept. mot. Skills*, **18**, 173–4.

EYSENCK, H. J. & MAXWELL, A. E. (1961). Reminiscence as a function of drive. *Brit. J. Psychol.* **52**, 43–52.

EYSENCK, H. J. & WILLETT, R. A. (1961). The measurement of motivation through the use of objective indices. *J. ment. Sci.* **107**, 961–8.

FELDMAN, M. P. (1963). A reconsideration of the extinction hypothesis of warm-up in motor behavior. *Psychol. Bull.* **60**, 452–9.

FELDMAN, M. P. (1964*a*). Pursuit rotor performance and reminiscence. In *Experiments in Motivation* (ed. H. J. EYSENCK). Oxford: Pergamon.

FELDMAN, M. P. (1964*b*). Drive and pursuit rotor reminiscence: the effect of an alien stimulus. In *Experiments in Motivation* (ed. H. J. EYSENCK). Oxford: Pergamon.

FLESCHER, P. (1941). L'amnesia retrograda dopo l'elettro shock. *Schweiz. Arch. Neurol. Psychiat.* **48**, 1–28.

GLEITMAN, H., MACHMIAS, J. & NEISSER, U. (1954). The S–R reinforcement theory of extinction. *Psychol. Rev.* **61**, 23–33.

GLICKMAN, S. E. (1958). Deficits in avoidance learning produced by stimulation of the ascending reticular formation. *Canad. J. Psychol.* **12**, 97–102.

GLICKMAN, S. E. (1961). Perseveration, neural processes and consolidation of the memory trace. *Psychol. Bull.* **58**, 218–33.

GRASSI, J. (1964). Experimental investigation of cortical inhibition in children. London: unpublished Ph.D. thesis, University of London Library.

GRICE, G. R. & REYNOLDS, B. (1952). Effects of varying amounts of rest on conventional and bilateral transfer 'reminiscence'. *J. exp. Psychol.* **44**, 247–52.

HAYES, K. J. (1953). Anoxia and convulsive amnesia in rats. *J. comp. physiol. Psychol.* **46**, 216–17.

HOLLAND, H. C. (1963). Massed practice and reactive inhibition, reminiscence and disinhibition in the spiral after-effect. *Brit. J. Psychol.* **54**, 261–72.

HOVLAND, C. I. (1951). Human learning and retention. In *Handbook of Experimental Psychology* (ed. S. S. STEVENS). London: Chapman and Hall.

HULL, C. L. (1943). *Principles of Behavior.* New York: Appleton-Century-Crofts.

IRION, A. L. (1949). Reminiscence in pursuit-rotor learning as a function of rest and of amount of pre-rest practice. *J. exp. Psychol.* **39**, 492–9.

JENSEN, A. R. (1961). On the reformulation of inhibition in Hull's system. *Psychol. Bull.* **58**, 274–98.

KENDRICK, D. C. (1960). The theory of 'conditioned inhibition' as an explanation of negative practice effects: an experimental analysis. In *Behaviour Therapy and the Neuroses* (ed. H. J. Eysenck). Oxford: Pergamon.

KIMBLE, G. A. (1949). An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.* **39**, 15–23.

KIMBLE, G. A. (1950). Evidence for the role of motivation in determining the amount of reminiscence in pursuit rotor learning. *J. exp. Psychol.* **40**, 248–53.

KIMBLE, G. A. & SHATEL, R. B. (1952). The relationship between two kinds of inhibition and the amount of practice. *J. exp. Psychol.* **44**, 355–9.

KOCH, S. (ed.) (1954). Clark versus Hull. In *Modern Learning Theory.* New York: Appleton-Century-Crofts.

LEUKEL, F. R. (1957). The effect of E.C.S. and pentothal anaesthesia on maze learning and retention. *J. comp. physiol. Psychol.* **50**, 300–6.

MCGAUGH, J. L. & PETRINOVICH, L. (1959). The effect of strychnine sulphate on maze-learning. *Amer. J. Psychol.* **72**, 99–102.

MCGAUGH, J. L. & PETRINOVICH, L. (1963). Comments concerning the basis of learning enhancement with central nervous system stimulants. *Psychol. Rep.* **12**, 211–14.

McGAUGH, J. L., WESTBROOK, W. H. & THOMSON, C. W. (1962). Facilitation of maze learning with post-trial injections of 5,7-diphenyl-1,3-diazadamantan-6-ol (1757 I.S.). *J. comp. physiol. Psychol.* **55**, 710–13.

McGEOCH, J. A. & IRION, A. L. (1952). *The Psychology of Human Learning.* New York: Longmans, Green and Co.

MÜLLER, C. E. & PILZECKER, A. (1900). Experimentelle Beiträge zur Lehre vom Gedächtnis. *Z. Psychol.* Suppl. 1.

OSGOOD, C. E. (1953). *Method and Theory in Experimental Psychology.* New York: Oxford University Press.

RACHMAN, S. (1962). Disinhibition and the reminiscence effect in a motor learning task. *Brit. J. Psychol.* **53**, 149–57.

RACHMAN, S. & GRASSI, J. (1965). Reminiscence, inhibition and consolidation. *Brit. J. Psychol.* **56**, 157–62.

RUSSELL, R. W. & HUNTER, W. S. (1937). The effects of inactivity produced by sodium amytal on the retention of the maze habit in the albino rat. *J. exp. Psychol.* **20**, 426–36.

RUSSELL, W. R. & NATHAN, P. W. (1946). Traumatic amnesia. *Brain*, **69**, 280–300.

SAVAGE, R. D. (1964). Electro-cerebral activity, extraversion and neuroticism. *Brit. J. Psychiat.* **110**, 98–100.

SHAGASS, C. & SCHWARTZ, M. (1963). Neurophysiological dysfunction associated with some psychiatric disorders. *Psychiat. Res. Rep.* **17**, 130–52.

SPIELMANN, I. (1963). The relation between personality and the frequency and duration of involuntary rest pauses during massed practice. London: unpublished Ph.D. thesis, University of London Library.

STAR, K. H. (1957). An experimental study of 'reactive inhibition' and its relation to certain personality traits. London: unpublished Ph.D. thesis, University of London Library.

STELLAR, E. (1957). Physiological psychology. *Annu. Rev. Psychol.* **8**, 415–36.

THIESSEN, D. D., SCHLESINGER, K. & CALHOUN, W. H. (1961). Better learning: neural enhancement or reduced interference? *Psychol. Rep.* **9**, 493–6.

THOMPSON, R. (1958). The effect of intracranial stimulation on memory in cats. *J. comp. physiol. Psychol.* **51**, 421–6.

THOMPSON, R. & PRYER, R. S. (1956). The effect of anoxia on the retention of a discrimination habit. *J. comp. physiol. Psychol.* **49**, 297–300.

TIBER, N. & KENNEDY, W. G. (1964). The effects of incentives on the intellligence test performance of different social groups. *J. consult. Psychol.* **28**, 187–8.

WALKER, E. L. (1958). Action decrement and its relation to learning. *Psychol. Rev.* **65**, 129–42.

WALKER, E. L. & TARTE, R. D. (1963). Memory storage as a function of arousal and time with homogeneous and heterogeneous lists. *J. verb. Learn. verb. Behav.* **2**, 113–19.

WASSERMAN, H. N. (1951). The effect of motivation and amount of pre-rest practice upon inhibitory potential in motor learning. *J. exp. Psychol.* **42**, 162–72.

WILLETT, R. A. (1964). Eyeblink conditioning and situation-induced anxiety drive. In *Experiments in Motivation* (ed. H. J. Eysenck). Oxford: Pergamon.

WILLETT, R. A. & EYSENCK, H. J. (1962). An experimental study of human motivation. *Life Sci.* no. 4, 119–27.

WILLIAMS, H. L., LUBIN, A. & GOODNOW, J. J. (1959). Impaired performance with acute sleep loss. *Psychol. Mon.* **73**, no. 14.

WILLIAMS, M. (1950). Memory studies in electric convulsion therapy. *J. Neurol. Neurosurg. Psychiat.* **13**, 30–5.

ZUBIN, J. & BARRERA, S. E. (1941). Effect of electric convulsive therapy on memory. *Proc. Soc. exp. Biol., N.Y.*, **48**, 596–7.

(*Manuscript received* 26 *November* 1964)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 183–195
Printed in Great Britain

EFFECTS OF AGE ON SHORT-TERM STORAGE AND SERIAL ROTE LEARNING

By JAMES INGLIS AND MARY N. ANKUS
Queen's University, Kingston, Ontario

Previous studies have shown that, as age advances, there is a progressive decrease in the ability to respond sequentially to simultaneous stimuli. It has been suggested that this impairment with age is due to a decline in the efficiency of some short-term storage process. In those studies, however, the relation of this process to performance on other learning tasks had not been traced in normal subjects, nor have the possible effects of changes in sensory acuity, perception or attention been adequately controlled.

This paper describes the results obtained from 120 normal subjects aged from 11 to 70 yr in their reproduction of dichotic digits (*a*) when the order of recall is left to the free choice of the subject, (*b*) when the order of recall has been specified before, and (*c*) after these digits have been delivered. The correlation of performance on this task with performance on serial learning is also described.

The results obtained confirm the view that it is a change in some short-term storage process rather than in any perceptual function which principally affects performance in this kind of experimental situation. Evidence was also obtained that short-term storage is an important, but not a unique component, of longer term learning.

Broadbent (1958) has suggested that the ability successively to reproduce a series of digits when this has been presented as two simultaneous half-sets to each ear, depends upon some short-term storage mechanism. Thus, in one experiment Broadbent (1957) showed that when digit-span stimuli were relayed at a speed of 2 per sec through headphones, one half of the span to one ear and, simultaneously, the other half of the span to the other ear, subjects tested under these conditions could reproduce the digits sequentially. Furthermore, in such reproduction, elements from one half of the span were rarely alternated with elements from the other half. The first half-set recalled, however, commonly contained fewer errors than the second half, producing a kind of serial-order effect. Broadbent has suggested two kinds of mechanism which may underlie such performance on this modified digit-span test: first, a '*p*-system' which can only pass information successively; secondly, an '*s*-system' which can store excess information arriving, for example, when the *p*-mechanism is already fully occupied in transmitting information from another channel. In each case the half-set of digits recalled first passes directly through the *p*-system while the half-set recalled second spends some time in storage.

There is a growing body of evidence that human memory may depend, in part at least, upon the efficiency of this short-term storage stage of the learning process. Inglis (1960), for example, has suggested that the defect of acquisition which has been found in elderly psychiatric patients suffering from memory disorder might be based upon a breakdown of this kind of storage mechanism. Such patients should therefore show disturbance of recall in those half-sets of digits reproduced second in the dichotic stimulation situation, relative to the performance of a matched group of patients without memory disorder. This expectation has been confirmed by Inglis & Sanderson (1961) for the case in which the two channels for simultaneous stimulation were the two ears. Caird & Inglis (1961) have confirmed these results and extended

them to the case in which the ear and eye were together presented with different digits.

Inglis (1962*a*) has also pointed out that since increasing age also seems to affect learning capacity, and as such impairment may also depend on changes in short-term storage, it might be that responses to dichotic stimulation would vary with normal ageing. If age primarily affects storage then the reproduction of the first half-set recalled should not be affected by advancing years. If the second half-set recalled must pass through the storage process then the recall of these digits should be affected by age.

This expectation has been confirmed in two studies. In the first of these Inglis & Caird (1963) examined the changes in sequential responses to simultaneous stimulation in 120 subjects between the ages of 11 and 70 yr. In the second study, by Mackay & Inglis (1963), the responses of 160 subjects between the ages of 11 and 90 yr were examined. The results of these two experiments, carried out on different subjects by different experimenters, proved to be in very close agreement. As age increases there is little or no significant impairment in the ability to recall the half-spans reproduced first. Progressively and significantly greater difficulty is, on the other hand, shown in the reproduction of the second half-spans as age advances. Furthermore, the longer the span to be recalled the greater the difference between the first and second half-spans.

In young patients, before and after dominant temporal lobectomy for the relief of temporal lobe epilepsy, Kimura (1961*a*, *b*) found defects in the recall of dichotically presented digits. She maintained that her results show the principal effects of such brain damage to be upon auditory *perceptual* functions. It has been argued by Inglis (1962*b*), however, that her data are not incompatible with the notion that a disturbance of auditory *storage* is the chief result of dominant temporal-lobe dysfunction. It is known, from the work of Meyer & Yates (1955), Milner (1958) and Meyer (1959) that damage in this brain area impairs the storage of other kinds of auditory material as, for example, is shown in paired-associate learning. Kimura (1962) has argued further for her view but has not presented an analysis of her data in a form which would permit an evaluation of the relative importance of the order-accuracy of recall as compared with the side-accuracy of recall. To ascertain whether or not Kimura's (1961*a*) data reflect a storage dysfunction similar to that shown in the senile and the aged it would be necessary to analyse her data in a different way from the one she chose. Account would have to be taken, not only of the total recall score, or simply of the side recalled, whether this be ipsilateral or contralateral to brain injury, but also of which half-span (i.e. the first or second recalled) is principally affected in each case. The model derived from Broadbent (1958) and advocated by Inglis (1962*b*) would predict that the main significant difference found in response to dichotic stimulation between left-sided (dominant) and right-sided (non-dominant) temporal lobe patients is due principally, if not entirely, to the poorer recall by the former group of the second half-spans, thus reflecting an impairment of storage. It should be noted here that Bryden (1963) has produced some evidence which tends to support Kimura's view. In one dichotic listening experiment Bryden found that normal students tended to reproduce material from the right ear more accurately than material from the left, and that this superiority of recall was not simply a function of the order of

report. Results from the experiments to be described below, however, cast some doubt upon the generality of these findings.

Evidence on the effect of maturation in childhood on successive responding to simultaneous stimulation has also been provided by Kimura (1963). Again she has shown in 145 normal children between the ages of 4 and 9 yr that the digits arriving at the right ear are more accurately reported than digits arriving at the left; this she interprets as meaning that those digits arriving at the ear opposite the dominant hemisphere are more efficiently *recognized*. Since no analysis is made, however, of order of report in this study either, the possibility remains that the right ear produced the more accurate report because the material delivered there was reproduced first in sequence and hence was subject to less storage decay.

It is certainly the case that students of M. S. Rabinovitch of McGill University have shown that children with learning defects (Witelson, 1962) but not simply with emotional disturbances (Litvack, 1963) also show a defect in reporting dichotic digits. Following Kimura and Bryden, however, emphasis was placed in these studies upon the total score achieved by the subjects rather than upon the possibility of an increased order-effect of the kind which would be expected to result from a storage defect. This led Witelson (1962), for example, to conclude that children with learning problems are handicapped by a defect of auditory perception. It might, however, seem more reasonable to expect them to show an impairment of auditory storage. It would certainly seem worthwhile to examine this possibility.

It has been pointed out by Kimura (1962) that it is not, of course, incumbent on those who use Broadbent's technique also to employ his theoretical model. This argument would be the more persuasive if a powerful alternative model were at the same time offered. It seems, however, that Broadbent's original hypothesis concerning storage can better comprehend the results of dichotic listening studies in cases of both generalized and specific learning defect and also generate more new and testable expectations than any hypothesis yet put forward to account for these phenomena in terms of perceptual processes. The investigations to be reported in this paper were designed with the intention of further exploring Broadbent's model in relation to changes in short-term storage with age.

Previous studies reported by Inglis & Caird (1963) and by Mackay & Inglis (1963) on the effects of age on sequential responses to simultaneous stimulation failed to take into account a number of important considerations. In the first place the subjects tested in these studies were not given any other kind of learning test. Thus the conclusion that there is, in normal subjects, some connexion between the age-related deficit in dichotic digit recall and other forms of decreased learning capacity (Jerome, 1959) is based only on inference and not on direct demonstration.

In the second place these studies did not ensure any systematic control over the effects of the ear to which responses were made, but analysed results only in terms of order of response, ignoring the laterality effects which Bryden (1963) found to be of importance in young normal subjects.

In the third place, it could be maintained that the impairment shown by elderly subjects in the recall of dichotic digits might be due to a failure to hear, or at least to a failure to attend to or to perceive the second half-spans on which they showed most errors of report. If this were the case it would support Kimura's interpretation that

defects of performance on this task reflect mainly upon accuracy of perception rather than upon adequacy of storage.

The present study comprises three main sections which were designed to deal with these considerations.

Method

Subjects

The subjects tested in this study were 120 persons between 11 and 70 years of age. None of these was known to be suffering from any gross mental or physical handicap; none had taken part in any previous studies. There were 20 people in each decade group: 10 male and 10 female. They were taken at random (but not, strictly speaking, 'randomly') from the population of Kingston, a city in Ontario of about 50,000 inhabitants.

Procedure

All the subjects were given the orthodox digit-span forward test described by Wechsler (1955) and a serial anticipation rote learning test of the kind used by Bromley (1958) in his study of the effects of age on learning. The score on the rote learning task was obtained as follows. Twelve nonsense syllables of 80 % association value (e.g. TAS, DIT, PEL) were taken from the lists provided by Hilgard (1951) and printed in large letters on separate cards. The experimenter exposed the syllables one by one at a rate of approximately one every 2 sec in serial order. After the first complete presentation the subject went through the list trying to anticipate each syllable, correcting himself or being corrected by the experimenter whenever he made a mistake. The score obtained was the number of correct anticipations on the eleventh trial. If a subject managed to learn the complete series accurately before the eleventh presentation he was given an extra credit for each unused trial.

Table 1. *Digits used for binaural stimulation*

Channel 1	Channel 2	Channel 1	Channel 2
Practice series			
3	Blank		
Blank	7		
3	7		
Test series		Test series	
5	8	5638	2941
7	6	9754	8362
4	1	6542	7918
6	3	9356	4271
39	72	81342	96571
85	17	74682	31579
38	59	57841	29356
65	28	38671	15429
592	174	251364	746982
793	462	984375	753162
479	836	451328	238691
584	719	438695	965127

The apparatus used to present the binaural stimuli was a 'Roberts 990' tape recorder. Different sets of digits were recorded on two channels as shown in Table 1. The test series were played at the rate of one digit every two-thirds of a second. There were four of each length of span with from 1 to 6 digits per half-span, making 24 spans in all for each condition.

The administration of this test material fell into the three following parts.

(i) *Free recall*. This will be labelled the '*F*' condition and was the one to which every subject was exposed immediately after the serial learning test described above.

Each subject was told, 'You are going to hear a number. Tell me what you hear.' The spoken digit 3 was then played on channel 1. If the subject reproduced this correctly, the same procedure was repeated for the digit 7 on channel 2. If the subject failed to respond or gave the

wrong number the stimulus was repeated and, if necessary, the volume increased until the correct response was made. Secondly, the subject was told, 'Now you are going to hear two numbers together, one in the right ear and one in the left ear. You will hear them both at the same time. Tell me what you hear.' The two channels then played the spoken digits 3 and 7. If the subject reproduced the correct digits (in either order 3–7 or 7–3) the experimenter went on to the test series; if not this item was repeated until both numbers were correctly reproduced. The foregoing provided a practice series which served to get subjects used to the experimental situation. Thirdly each subject was told, 'Now you are going to hear n numbers, $n/2$ in each ear' (where n was 2, 4, 6, 8, 10, or 12). 'Tell me what numbers you hear.'

Responses were scored as follows: the first digit repeated determined in each case which channel was taken to be the half-span recalled first. The score obtained was the average number of correct responses for each half-set of digits, taking each digit's position in the series into account. If, as sometimes happened, the first digit recalled was not in fact the first term of either half-series then this sequence was regarded as 'spoiled' and excluded from the scoring calculations. This procedure was used in order to provide a further repetition of the experiments carried out on similar groups by Inglis & Caird (1963) and Mackay & Inglis (1963).

It can be seen that no control was exercised by these instructions over the laterality of the order of recall.

(ii) *Recall order specified before delivery of digits*. This will be labelled the '*B*' condition and half of the subjects were given the test material under this condition directly after the '*F*' condition. The remaining subjects were tested in this way after they had also experienced the '*A*' condition defined below.

The side to be recalled first was indicated by means of a signal panel set in front of the subject simply bearing the printed words 'right' and 'left'; a red light was switched on above the appropriate word by the experimenter, in this case immediately *before* each span of the dichotic digits was delivered. The experimenter gave each subject the following instructions. 'Now we will go through the same procedure again but this time *before* you hear the numbers I will signal in which ear order you are to tell me the numbers. For example, if I signal "left" tell me first the numbers you hear in your left ear and then what you hear in your right ear. If I signal "right" tell me first the numbers you hear in your right ear.'

In this series there were 8 sets of each length of span. The subject was signalled to recall the left ear first on half the occasions and the right ear on the other occasions of the delivery of a given length of span. The right and left signals within each test series were given in random order. Thus there were four of each length of span from 1 to 6 digits per half-span, making 24 spans in all which required recall from the left ear first (*B–L*) and 24 spans which required the right ear to be recalled first (*B–R*).

An analogous method of administration had previously been employed by Broadbent (1957). It was used in the present experiment to evaluate the notion that the changes with age which had previously been shown for this task might be due, for example, to an age-related hearing loss in one ear as compared to the other. If it were the case that as people grow older they tend to suffer such a differential loss in auditory acuity then the apparent deficit in the ear recalled second under conditions of free response might be explained by this kind of peripheral difficulty. If, however, an increased order effect were to appear with age regardless of ear-specification then this hypothesis would not be supported.

(iii) *Recall order specified after delivery of digits*. This will be labelled the '*A*' condition: half of the subjects were given the test material under this condition directly after the '*F*' condition. The remaining subjects were tested in this way after they had experienced first the '*F*' and then the '*B*' condition defined above.

The side to be recalled first was again specified by means of the light signals, in this case immediately *after* the dichotic digits were delivered. The experimenter gave each subject the following instructions. 'Now we will go through the same procedure again but this time immediately *after* you hear the numbers I will signal in which ear order you are to tell me the numbers. For example, if I signal "left" tell me first the numbers you heard in your left ear and then what you heard in your right ear. If I signal "right" tell me first the numbers you heard in your right ear.'

There were again 8 sets of each length of span. The right and left signals also appeared an equal number of times in random order. There were, therefore 24 spans requiring recall from the left ear first (*A–L*) and 24 requiring recall from the right ear first (*A–R*).

This method had also previously been used by Broadbent (1957). It was employed in the present experiment to control for the possible effects of differences in attention. It might be argued, for example, that previous findings were due to the fact that as people grow older they will only pay attention to the material coming into one ear in a dichotic digit experiment. It can be seen that the '*B*' condition does not provide any control over this possibility. If, however, the subjects only know after the digits have been delivered which side they have to report first and, further, if they only attend to one side, then the ear order effect should vanish, or at least be greatly diminished since they cannot know beforehand to which ear they should have paid most attention.

Results and discussion

The results of this investigation may again be most conveniently considered under three heads, as follows:

(i) *Free recall and rote learning*

The differences between the means of the six age groups for orthodox digit-span performance, serial rote learning and for the first and second half-sets reproduced under the free recall condition of dichotic stimulation were assessed by trend analyses of variance (Edwards, 1960) as shown in Table 2.

The results for the various digit span tests are in close accord with previous findings (Inglis & Caird, 1963; Mackay & Inglis, 1963). The orthodox digit span forward test shows no significant change with age. There is only one case in six in which the first half-set recalled in the dichotic condition shows any change with age. Conversely, there is only one case in six in which the second half-span fails to show a significant association with age. The trend analyses thus indicate that the accuracy of recall of the second half-sets decreases with age in accordance with expectation and previous findings.

These results very clearly demonstrate the effects of increasing age on the dichotic digit half-spans of different lengths. As age increases there is little significant impairment in ability to recall the first half-spans. Progressively greater difficulty is shown, however, in the reproduction of the second half-spans. Furthermore, the longer the span to be recalled the greater the difference between the first and second half-spans.

On the basis of these data an attempt was made to evaluate Bryden's (1963) finding that the right ear is predominantly chosen as the first ear under conditions of free recall. For each decade group there were 20 subjects who responded 24 times: this gives a possible total of 480 choices between recalling the right or left ear first; totals less than this figure reflect the number of 'spoiled' responses. The actual proportions of choice for these groups is shown in Table 3.

The right-left differences in each group were assessed by means of a sign test described by Siegel (1956). While four out of six of these differences are significant they are not, in fact, very systematic as between groups. The right ear is chosen predominantly as the ear to be first recalled by the 11–20's and the 61–70's; the left ear is chosen by the 21–30's and the 51–60's. The number of left-handed people (five out of a total of 120) in any decade group was very small, which makes it unlikely that these findings are due to any differences, say, in cerebral dominance between the groups. These results, then, already suggest that order is more crucial than laterality in producing impairment of recall.

Table 2. *Means, standard deviations and the significance of the differences between the means of the six age groups for digit-span, serial rote learning and dichotic digits ('F' condition)*

		Age groups												Tests of significance					
		11–20		21–30		31–40		41–50		51–60		61–70		Over-all		Linear component		Quadratic component	
Variables		*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*F*	*P*	*F*	*P*	*F*	*P*
Age		16·75	2·51	24·65	3·08	36·35	2·48	44·70	3·17	54·60	2·91	66·10	2·89	—	—	—	—	—	—
Digit-span forward		6·70	1·18	6·65	1·43	6·30	1·38	6·20	0·90	5·85	0·99	6·35	0·88	1·529	NS	—	—	—	—
Learning score		11·90	3·37	10·00	4·55	7·20	2·02	7·35	3·18	5·95	2·54	4·20	2·22	16·112	< 0·005	76·058	< 0·005	1·002	NS
No. digits per ½-span	Recall channel																		
1	1st	1·00	0·00	0·98	0·02	1·00	0·00	1·00	0·00	0·98	0·10	0·98	0·10	1·333	NS	—	—	—	—
	2nd	1·00	0·00	0·96	0·02	0·95	0·10	0·86	0·17	0·86	0·14	0·66	0·35	8·706	< 0·005	36·176	< 0·005	4·059	< 0·05
2	1st	1·61	0·37	1·59	0·41	1·45	0·36	1·45	0·42	1·53	0·51	1·68	0·36	0·496	NS	—	—	—	—
	2nd	1·30	0·44	1·34	0·67	0·93	0·58	0·96	0·64	0·90	0·75	0·75	0·60	2·909	< 0·025	12·086	< 0·005	0·133	NS
3	1st	2·05	0·74	2·05	0·77	1·79	0·76	1·83	0·73	1·49	0·58	2·10	0·63	2·194	NS	—	—	—	—
	2nd	1·26	1·00	1·34	1·21	0·90	1·04	0·69	0·88	0·48	0·66	0·60	0·89	2·736	< 0·025	11·491	< 0·005	0·339	NS
4	1st	2·61	1·32	2·26	1·14	2·23	1·21	1·98	0·81	1·98	1·10	2·63	1·21	1·323	NS	—	—	—	—
	2nd	1·09	0·98	1·04	1·30	0·51	0·84	0·34	0·57	0·14	0·20	0·14	0·41	5·645	< 0·005	25·675	< 0·005	0·884	NS
5	1st	2·68	1·55	1·88	0·79	2·01	0·99	2·08	1·16	2·00	1·12	2·86	1·41	2·363	< 0·05	0·377	NS	1·050	NS
	2nd	0·41	0·71	0·14	0·25	0·06	0·14	0·01	0·02	0·10	0·20	0·09	0·17	3·706	< 0·005	8·376	< 0·005	8·413	< 0·005
6	1st	2·30	1·59	2·21	1·64	2·80	1·59	2·24	1·04	2·25	1·69	2·90	1·61	0·780	NS	—	—	—	—
	2nd	0·15	0·26	0·09	0·23	0·05	0·10	0·04	0·03	0·03	0·10	0·00	0·00	2·240	NS	—	—	—	—

The results for the rote learning test are shown graphically in Fig. 1. The age decrement for this test is sharper than it was in Bromley's (1958) study, but this may be due to the fact that he only tested intellectually superior subjects.

Table 3. *Right and left responses showing ear preference under conditions of free recall with level of significance assessed by means of a sign test*

Age groups	No. of right responses	No. of left responses	Level of significance
11–20	249	181	< 0·001
21–30	160	226	< 0·001
31–40	216	221	NS
41–50	215	230	NS
51–60	156	257	< 0·001
61–70	246	199	0·01
Total	1242	1354	< 0·01

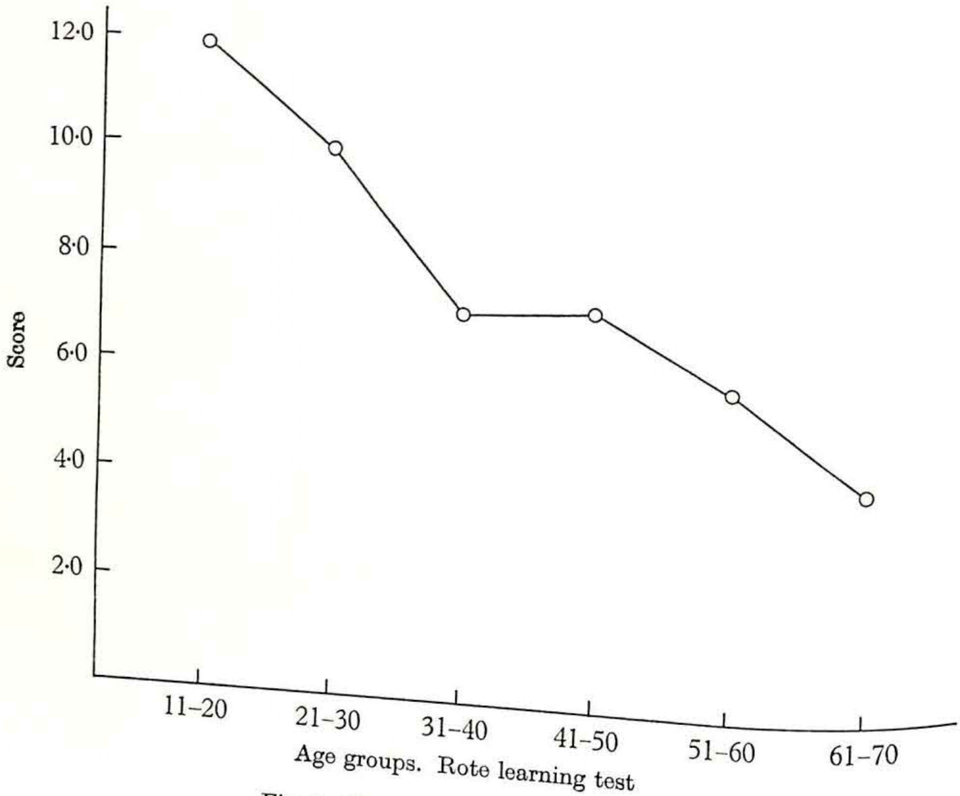


Fig. 1. Scores on serial learning test.

In order to investigate the relation of performance on the dichotic digits to performance on the rote learning test it was necessary to make some arbitrary choice of many possible associations. Since the notion of a short-term storage process had been based by Broadbent (1958) mainly on the study of the performance of young normal individuals on the second half-span of series with three digits in each half-span, it was decided to use this item as the critical aspect of dichotic digit performance. An analysis of covariance was then carried out (McNemar, 1949) in which the longer term memory represented by the rote learning test was regarded as the dependent variable. This analysis is shown in Table 4.

These results show that while there is a considerable relation between the two measures (i.e. between group correlation of 0·89, $P < 0{\cdot}01$) there is still an effect of age on rote learning after the influence of short-term storage has been partialled out. This may be interpreted to mean that the impairment of short-term storage with age is an important but not unique component of the decrease in learning capacity associated with advancing years.

Table 4. *Analysis of covariance on a supplementary variable* (X) *and a dependent variable* (Y) *where X represents short-term memory and Y serial rote learning*

	Total	Within	Between
Sum of products	197·31	109·19	88·12
Sum of squares for X	118·63	105·92	12·71
Sum of squares for Y	1863·47	1091·70	771·77
Degrees of freedom	119	114	5
Correlation	0·42	0·32	0·89
Degrees of freedom	118	113	4
Adjusted sum of squares for Y	1535·30	979·14	556·16
Degrees of freedom	118	113	5
Mean square	—	8·66	111·23

$F = 111{\cdot}23/8{\cdot}66 = 12{\cdot}84$; D.F. 5 and 11; $P < 0{\cdot}001$.

(ii) *Recall order specified before delivery of digits*

The differences between the groups on the first and second half-spans of the dichotic digits were again assessed separately by means of trend analyses. It will be recalled that there were two subclasses in this section, *B–L* and *B–R*. The analysis of the former set of results is shown in Table 5.

In this case the effect of age is again principally on the second half-sets although in the case of 2 digits per half-span the first half-set also appears to be affected by age. The resemblance of these results to those secured under conditions of free recall is striking.

A statistical analysis of the *B–R* condition is, on the face of it, rather less clear-cut, as shown in Table 6. This shows an age-related change in four of the six first half-sets as well as in five of the six second half-sets; but the order effect still exists and the age-related decline is more marked in the digits recalled second.

When the results from the two conditions *B–L* and *B–R* are considered together they refute the notion that, for example, any unilateral decrease in auditory acuity could account for the appearance of the order effect noted under conditions of free recall. A marked age-related order effect appears which is independent of laterality.

(iii) *Recall order specified after delivery of digits*

The results of trend analyses of variance for the *A–L* condition are shown in Table 7. There is again a striking resemblance to the results secured under conditions of free recall. Similar analyses of the data obtained in the *A–R* condition are shown in Table 8. Here again, although in five out of six cases performance on the second half-spans varies with age, so does performance on two of the six first half-spans. It can be seen that the age-related order effect is clearly shown; the decline with age in the half-sets recalled second being more profound than in the half-sets recalled first.

Table 5. *Means, standard deviations and the significance of the differences between the means of the six age groups in the dichotic digits (B–L condition)*

		Age groups												Tests of significance					
		11–20		21–30		31–40		41–50		51–60		61–70		Over-all		Linear component		Quadratic component	
No. digits per ½-span	Recall channel	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*F*	*P*	*F*	*P*	*F*	*P*
1	1st	1·00	0·00	1·00	0·00	1·00	0·00	1·00	0·00	0·98	0·10	0·94	0·29	1·200	NS	—	—	—	—
	2nd	1·00	0·00	1·00	0·00	0·95	0·14	0·98	0·10	0·95	0·10	0·79	0·32	5·478	< 0·005	17·870	< 0·005	5·217	< 0·025
2	1st	1·91	0·22	1·92	0·17	1·85	0·17	1·85	0·32	1·70	0·36	1·66	0·47	3·185	< 0·025	13·123	< 0·005	1·000	NS
	2nd	1·88	0·26	1·90	0·25	1·78	0·41	1·68	0·42	1·44	0·71	1·25	0·80	5·610	< 0·005	25·142	< 0·005	2·326	NS
3	1st	2·67	0·53	2·66	0·49	2·60	0·44	2·32	0·74	2·38	0·67	2·40	0·78	1·269	NS	—	—	—	—
	2nd	2·08	0·87	1·75	1·12	1·28	1·09	1·20	1·11	1·04	0·90	0·80	0·97	4·315	< 0·005	20·463	< 0·005	0·632	NS
4	1st	3·24	0·95	3·43	0·59	3·09	0·96	3·34	0·85	2·95	0·91	2·90	0·99	1·149	NS	—	—	—	—
	2nd	1·33	1·27	1·30	1·23	0·60	1·03	0·32	0·64	0·32	0·51	0·40	0·79	5·050	< 0·005	19·451	< 0·005	2·937	NS
5	1st	3·73	1·41	3·15	1·35	3·01	1·07	3·11	1·13	2·87	1·50	3·05	1·47	1·013	NS	—	—	—	—
	2nd	0·74	1·23	0·35	0·51	0·26	0·26	0·15	0·30	0·09	0·20	0·12	0·14	3·553	< 0·005	13·785	< 0·005	3·417	NS
6	1st	3·88	1·70	3·46	1·55	3·43	1·67	3·73	1·39	2·76	1·68	2·97	1·40	1·518	NS	—	—	—	—
	2nd	0·22	0·33	0·14	0·36	0·04	0·17	0·04	0·14	0·00	0·00	0·04	0·17	2·566	< 0·05	9·434	< 0·005	3·151	NS

Table 6. *Means, standard deviations and the significance of the differences between the means of the six age groups in the dichotic digits (B–R condition)*

		Age groups												Tests of significance					
		11–20		21–30		31–40		41–50		51–60		61–70		Over-all		Linear component		Quadratic component	
No. digits per ½-span	Recall channel	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*m*	*sd*	*F*	*P*	*F*	*P*	*F*	*P*
1	1st	1·00	0·00	0·99	0·00	1·00	0·00	1·00	0·00	0·80	0·41	0·85	0·36	3·176	< 0·025	9·647	< 0·005	1·353	NS
	2nd	1·00	0·00	0·99	0·00	1·00	0·00	1·00	0·00	0·78	0·40	0·85	0·36	3·880	< 0·005	11·000	< 0·005	1·260	NS
2	1st	1·88	0·32	1·96	0·10	1·86	0·26	1·89	0·24	1·55	0·74	1·41	0·80	4·220	< 0·005	15·106	< 0·005	4·403	< 0·05
	2nd	1·70	0·44	1·83	0·30	1·59	0·64	1·53	0·46	0·93	0·24	0·95	0·89	8·021	< 0·005	32·396	< 0·005	2·505	NS
3	1st	2·63	0·70	2·78	0·28	2·62	0·51	2·62	0·77	1·83	1·21	2·24	0·97	4·550	< 0·005	11·863	< 0·005	0·685	NS
	2nd	2·20	0·89	1·85	0·99	1·80	1·00	1·40	0·94	0·82	0·88	0·79	0·93	7·876	< 0·005	37·230	< 0·005	0·080	NS
4	1st	3·32	1·11	3·31	0·94	3·16	1·01	3·32	1·07	2·18	1·58	2·26	1·39	4·191	< 0·005	14·405	< 0·005	1·999	NS
	2nd	1·63	1·45	1·10	1·20	0·80	1·18	0·42	0·80	0·33	0·85	0·15	0·27	5·816	< 0·005	27·379	< 0·005	1·517	NS
5	1st	3·82	1·21	3·53	1·23	3·48	1·43	3·48	1·10	2·58	1·77	2·84	1·82	2·147	NS	—	—	—	—
	2nd	0·52	0·57	0·24	0·39	0·10	0·24	0·09	0·17	0·03	0·10	0·03	0·10	7·443	< 0·005	28·485	< 0·005	7·412	< 0·025
6	1st	3·53	1·71	3·36	1·79	3·50	1·61	3·63	1·42	2·42	2·19	2·58	1·53	1·886	NS	—	—	—	—
	2nd	0·23	0·40	0·12	0·20	0·10	0·20	0·10	0·20	0·05	0·14	0·05	0·10	1·909	NS	—	—	—	—

Table 7. *Means, standard deviations and the significance of the differences between the means of the six age groups in the dichotic digits (A–L condition)*

No. digits per ½-span	Recall channel	11–20 m	11–20 sd	21–30 m	21–30 sd	31–40 m	31–40 sd	41–50 m	41–50 sd	51–60 m	51–60 sd	61–70 m	61–70 sd	Over-all F	Over-all P	Linear component F	Linear component P	Quadratic component F	Quadratic component P
1	1st	1·00	0·00	1·00	0·00	1·00	0·00	1·00	0·00	0·95	0·22	0·94	0·22	1·059	NS	—	—	—	—
	2nd	0·99	0·00	0·98	0·10	1·00	0·00	0·98	0·10	0·91	0·24	0·76	0·33	5·125	< 0·005	22·500	< 0·005	8·156	< 0·025
2	1st	1·75	0·37	1·91	0·17	1·74	0·52	1·63	0·64	1·60	0·59	1·57	0·50	2·101	NS	—	—	—	—
	2nd	1·60	0·58	1·73	0·39	1·58	0·67	1·41	0·76	1·37	0·74	1·04	0·76	2·688	< 0·025	10·837	< 0·005	1·773	NS
3	1st	2·17	0·79	2·28	0·81	2·17	0·90	2·09	0·74	1·75	0·93	2·15	0·98	0·922	NS	—	—	—	—
	2nd	1·72	1·09	1·86	1·01	1·42	1·08	1·09	1·08	0·70	1·08	0·81	0·91	4·214	< 0·005	18·398	< 0·005	0·000	NS
4	1st	2·48	1·20	2·61	1·11	2·24	1·18	2·03	1·32	1·86	1·39	2·70	1·16	1·491	NS	—	—	—	—
	2nd	1·24	1·18	1·05	1·05	0·60	1·06	0·10	0·24	0·26	0·56	0·44	0·86	5·085	< 0·005	17·048	< 0·005	5·482	< 0·025
5	1st	2·48	1·59	1·65	1·23	1·98	1·31	1·77	1·38	1·62	1·45	2·36	1·48	1·337	NS	—	—	—	—
	2nd	0·67	0·98	0·40	0·65	0·16	0·44	0·12	0·24	0·12	0·40	0·04	0·14	3·759	< 0·005	15·338	< 0·005	2·910	NS
6	1st	2·33	1·67	1·37	1·15	1·72	1·28	1·32	1·56	1·95	1·66	2·05	1·18	1·577	NS	—	—	—	—
	2nd	0·18	0·32	0·08	0·25	0·05	0·10	0·03	0·10	0·03	0·10	0·08	0·17	1·778	NS	—	—	—	—

Table 8. *Means, standard deviations and the significance of the differences between the means of the six age groups in the dichotic digits (A–R condition)*

No. digits per ½-span	Recall channel	11–20 m	11–20 sd	21–30 m	21–30 sd	31–40 m	31–40 sd	41–50 m	41–50 sd	51–60 m	51–60 sd	61–70 m	61–70 sd	Over-all F	Over-all P	Linear component F	Linear component P	Quadratic component F	Quadratic component P
1	1st	1·00	0·00	1·00	0·00	0·95	0·22	0·94	0·22	0·78	0·41	0·83	0·37	2·441	< 0·05	9·926	< 0·005	0·147	NS
	2nd	0·98	0·10	1·00	0·00	0·95	0·22	0·93	0·22	0·74	0·41	0·69	0·41	4·613	< 0·005	19·427	< 0·005	2·373	NS
2	1st	1·82	0·36	1·93	0·17	1·60	0·82	1·63	0·56	1·66	0·63	1·47	0·67	1·642	NS	—	—	—	—
	2nd	1·64	0·48	1·83	0·28	1·39	0·84	1·20	0·78	0·78	0·86	0·85	0·83	6·984	< 0·005	29·962	< 0·005	0·117	NS
3	1st	2·44	0·76	2·46	0·65	1·95	0·94	1·96	0·97	1·80	1·14	1·53	0·98	3·361	< 0·005	15·344	< 0·005	0·002	NS
	2nd	1·93	0·98	1·87	1·08	1·41	1·21	1·03	0·93	0·95	1·02	0·79	0·95	4·584	< 0·005	21·580	< 0·005	0·228	NS
4	1st	2·43	1·00	2·32	1·09	1·92	1·39	2·13	1·21	1·58	1·24	1·45	1·34	2·107	NS	—	—	—	—
	2nd	1·30	1·33	1·15	0·95	0·44	0·72	0·32	0·55	0·27	0·49	0·20	0·57	6·914	< 0·005	29·022	< 0·005	3·223	NS
5	1st	2·27	1·38	1·49	0·91	1·91	1·10	2·26	1·17	1·85	1·70	1·75	1·70	0·987	NS	—	—	—	—
	2nd	0·25	0·49	0·12	0·25	0·09	0·24	0·08	0·30	0·06	0·20	0·11	0·35	0·882	NS	—	—	—	—
6	1st	1·83	1·17	1·89	1·20	1·39	0·88	2·08	1·52	1·71	2·07	1·65	1·72	0·513	NS	—	—	—	—
	2nd	0·35	0·68	0·08	0·17	0·13	0·26	0·03	0·10	0·03	0·10	0·04	0·14	3·122	< 0·025	9·204	< 0·005	3·724	NS

When the results from conditions A–L and A–R are taken together they suggest that the order-effect cannot be entirely explained by the notion that older subjects only attend to the material presented to one ear. If this were the case no increased order effect with age would be expected to appear in conditions A–L or A–R since the subjects could not, under these circumstances, anticipate which half-sets they would later be required to recall first.

It may be said, therefore, that the results of this experiment confirm, once again, that the responses made to dichotic stimulation under conditions of free recall are affected by age. Since it is the order effect which is exaggerated in older subjects this finding lends support to the view that some short-term storage mechanism is affected by advancing years. The impairment of this system is related to, but evidently not entirely responsible for, age-related defects in longer term learning.

The results obtained in this experiment when the order of recall was specified *before* delivery of dichotic digits are similar to the results from free recall. Since the age-related order effect still appears regardless of which side is specified for recall first, this finding is quite incompatible with the notion that changes in responses to dichotic listening with age might be due, for example, to some unilateral hearing loss which gets worse as people grow older.

The results obtained when order of report was specified *after* the delivery of the digits are also very similar to the results obtained under conditions of free recall. This outcome runs completely counter to the view that the changes observed with age might be due, for example, to the fact that older individuals only attend to, or manage to perceive, the material from a single channel under conditions of simultaneous stimulation.

These results give no support to Bryden's (1963) finding that report from the right ear is more accurate regardless of recall order. They also cast doubt upon Kimura's (1961*a*, 1963) conclusion that the advantage of right ear recall, when it occurs, is due to the superiority of perception on this side.

Neither, however, are the results of the third part of the experiment completely compatible with the version of Broadbent's hypothesis originally proposed. This would require some of the digits to have gone directly through the p-system while the others were being held in the s-system before the side to be recalled first was specified. Again a much diminished order-effect would be anticipated. The alternative notions put forward by Yntema & Trask (1963) may provide a more plausible alternative. They assume that both members of the dichotic pair are perceived and held in short-term storage at the time of presentation. Where no better criterion for classification exists, each digit is 'tagged' according to the side on which it was heard. The material reproduced second is recalled more poorly, not because it has passed through a mechanism different in kind from the material recalled first, but because it has simply been held longer in a similar system and been more subject to both trace decay and response interference. Inglis (1965) has suggested that such trace decay may be accelerated with age, and that age also increases sensitivity to interference. On this view it would be anticipated that as age advances the greatest impairment would be in the half-spans recalled second, just as in fact is shown by the experiments cited above.

The work described in this paper was supported by Grant no. 25 of the Ontario Mental Health Foundation whose help is most gratefully acknowledged. A brief version of the paper was presented at a Symposium on Mechanisms and Disturbances of Memory sponsored by Division 20 of the American Psychological Association at the Annual Meeting of the APA, Los Angeles, September 1964.

References

BROADBENT, D. E. (1957). Immediate memory and simultaneous stimuli. *Quart. J. exp. Psychol.* **9**, 1–11.

BROADBENT, D. E. (1958). *Perception and Communication.* Oxford: Pergamon.

BROMLEY, D. B. (1958). Some effects of age on short-term learning and remembering. *J. Gerontol.* **13**, 398–406.

BRYDEN, M. P. (1963). Ear preference in auditory perception. *J. exp. Psychol.* **65**, 103–5.

CAIRD, W. K. & INGLIS, J. (1961). The short-term storage of auditory and visual two-channel digits by elderly patients with memory disorder. *J. ment. Sci.* **107**, 1062–9.

EDWARDS, A. L. (1960). *Experimental Design in Psychological Research.* New York: Holt, Rinehart and Winston.

HILGARD, E. R. (1951). Methods and procedures in the study of learning. In *Handbook of Experimental Psychology* (ed. S. S. Stevens). New York: Wiley.

INGLIS, J. (1960). Dichotic stimulation and memory disorder. *Nature, Lond.*, **186**, 181–2.

INGLIS, J. (1962*a*). Effect of age on responses to dichotic stimulation. *Nature, Lond.*, **194**, 1101.

INGLIS, J. (1962*b*). Dichotic stimulation, temporal-lobe damage, and the perception and storage of auditory stimuli—a note on Kimura's findings. *Canad. J. Psychol.* **16**, 11–17.

INGLIS, J. (1965). Immediate memory, age and brain function. In *Behavior, Aging and the Nervous System* (eds. A. T. Welford and J. E. Birren). Springfield, Ill.: Thomas.

INGLIS, J. & CAIRD, W. K. (1963). Age differences in successive responses to simultaneous stimulation. *Canad. J. Psychol.* **17**, 98–105.

INGLIS, J. & SANDERSON, R. E. (1961). Successive responses to simultaneous stimulation in elderly patients with memory disorder. *J. abnorm. soc. Psychol.* **62**, 709–12.

JEROME, E. A. (1959). Age and learning—experimental studies. In *Handbook of Aging and the individual* (ed. Birren, J. E.). Chicago: University Chicago Press.

KIMURA, D. (1961*a*). Some effects of temporal-lobe damage on auditory perception. *Canad. J. Psychol.* **15**, 156–65.

KIMURA, D. (1961*b*). Cerebral dominance and the perception of verbal stimuli. *Canad. J. Psychol.* **15**, 166–71.

KIMURA, D. (1962). Perceptual and memory functions of the left temporal lobe—a reply to Dr Inglis. *Canad. J. Psychol.* **16**, 18–21.

KIMURA, D. (1963). Speech lateralization in young children as determined by an auditory test. *J. comp. physiol. Psychol.* **56**, 899–902.

LITVACK, B. M. (1963). Dichotic listening in emotionally disturbed children. Unpublished manuscript, McGill University.

MACKAY, H. A. & INGLIS, J. (1963). The effect of age on a short-term auditory storage process. *Gerontologia*, **8**, 193–200.

McNEMAR, Q. (1949). *Psychological Statistics.* New York: Wiley.

MEYER, V. (1959). Cognitive changes following temporal lobectomy for relief of temporal lobe epilepsy. *Arch. Neurol. Psychiat.* **81**, 299–309.

MEYER, V. & YATES, A. J. (1955). Intellectual changes following temporal lobectomy for psychomotor epilepsy. *J. Neurol. Neurosurg. Psychiat.* **18**, 44–52.

MILNER, B. (1958). Psychological defects produced by temporal lobe excision. *Publ. Ass. Res. nerv. ment. Dis.* **36**, 244–57.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw Hill.

WECHSLER, D. (1955). *Wechsler Adult Intelligence Scale Manual.* New York: Psychological Corp.

WITELSON, S. (1962). Perception of auditory stimuli in children with learning problems. Unpublished MSc. dissertation: McGill University.

YNTEMA, D. B. & TRASK, F. P. (1963). Recall as a search process. *J. verbal Learn. verbal Behav.* **2**, 65–74.

(*Manuscript received* 24 *September* 1964)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 197–215
Printed in Great Britain

MEMORY AND THOUGHT IN HUMAN INTELLECTUAL PERFORMANCE

By MICHAEL I. POSNER
*Department of Psychology, University of Wisconsin**

This paper is a review of efforts to extend the use of information techniques to tasks which are intellectual in nature. Complex tasks such as problem solving and concept formation are viewed in terms of simpler processes of information transformations and immediate memory. The first section of the paper considers efforts to describe the difficulty of transformations such as occur in arithmetic operations and concept utilization in terms of their informational parameters. The second part considers the relationship of these transformations to tasks which require retention. The final section extends the analysis to the complex sequential tasks of induction, problem solving and reading. The paper as a whole may be considered as a quantitative extension of the view of thinking as skilled performance (Bartlett, 1958).

I. Introduction

A large number of human tasks involve the presentation of information over time. Psychologists have abstracted from sequential tasks two pure modes by which man handles a stream of incoming information (Broadbent, 1958). The first of these is a relatively independent reponse to each event as, for example, in the reaction time experiment. This mode places maximal emphasis on transformation of each stimulus into an appropriate response but requires little retention. The second mode requires the subject to store the incoming stimuli and to reproduce them later (Posner, 1963). The memory span experiment represents an attempt to study this process.

Most tasks involve a combination of these pure modes. Reading a book, listening to a lecture, forming a concept or solving a problem are examples of complex tasks which involve the combination of these basic modes of processing. Each of these tasks requires the subject to take in information, to transform it in various ways by means of selection, classification and combination, and to store the product of these operations. Comprehension of a book or a lecture is not recall, producing the sum of a set of numbers does not require remembering the components, though without storage both tasks are clearly impossible. It is these serial tasks which place emphasis upon the storage of transformed information that seem to be most representative of intellectual performance.

Types of transforms

It is useful to consider a system which recognizes three types of transformations according to their informational properties. This taxonomy is based upon the relation between input and output information required for perfect performance of the task. First, information conservation tasks are those in which the subject is required to preserve all of the input information in his response. It is these tasks which have been studied most extensively by use of information measures. Both the standard reaction time and memory span experiments are conservation tasks in which any increase or decrease in information during transmission represents error. It is clear,

* Now at University of Oregon, Eugene, Oregon.

however, that man is more than just an information transmitting channel. He can act as a source of new information not present in a given stimulus or he can decrease or condense information, not merely through the making of errors, but also through a recoding which is a reflection of the stimulus information in a condensed output. A second type of transformation involves information creation. In this situation output information must exceed input if the subject is to perform the task. The classical experiment which fits this category is word association in which a single stimulus leads to a train of responses which exceeds the information value of the input. While this type of transformation has received some attention (Morin & Forrin, 1963; Shepard, 1963) the quantitative analysis of creation tasks is still relatively unexplored. The work discussed in this paper will lie in the third logical category of transformations; those which involve information reduction. Tasks such as addition, classification and selection are examples of information reducing transformations. Such transformations require the subject to produce a subset of the stimulus input. Any task in which the subject is required to map more than one stimulus point into a single response is a reduction task. The loss of information in these tasks clearly does not represent error, but rather is necessary to produce the required output.

Relation to Bartlett's theory

This type of framework for a study of human information processing is not entirely new. A number of investigators have proposed closely related taxonomies (Garner, 1962; Hunt, 1963). Of particular interest is the close relationship of this taxonomy and Bartlett's theory of thinking which itself arises out of experimental results in the area of skilled performance. Bartlett defines the types of thinking which he deals with in his system as 'the extension of evidence in accord with that evidence so as to fill up gaps in the evidence; and this is done by moving through a succession of interconnected steps which must be stated at the time, or left until later to be stated' (p. 75). In his system, Bartlett outlines three types of thinking which are analogous to the three types of informational transforms discussed above. First, he suggests that there is thinking within a closed system in which the response is in some way implicit in the stimulus input. This is clearly the case in the information reducing transforms. Within the closed system he discusses a type of mental activity which is called translation and in which there is a one to one correspondence between two coding schemes; clearly an analogue of information conservation. Finally, he suggests there is an open system in which the subject uses the evidence available in the stimulus to leap beyond the input and provide a new and creative solution. This is quite similar to the information creation notion. The scope of Bartlett's classification scheme is very large. The use of an informational analysis seems to capture many aspects of the verbal model and has the virtue of allowing some quantitative description.

The purpose of this paper is to examine complex intellectual tasks such as reading, induction, and problem solving in terms of the pure processing modes of transformation and retention. Accordingly, the next section considers a number of simple information reducing transformations in an attempt to specify the properties which govern the difficulty of such transforms. This section seeks to show the possibility of predicting the difficulty of such transforms from their informational properties. In

Section III the relation of such transformation processes to retention within simple serial tasks is examined. A number of experiments showing the trade off between transformation and retention are discussed. The final section discusses the application of these principles to relatively complex intellectual tasks.

II. Information reduction and thought

Several investigators have begun to apply the mathematics of information measures to predict the difficulty of operations in tasks which involve information reduction. Posner (1962, 1964*b*) hypothesized that the difficulty or amount of mental processing required in an information reduction task is directly related to the amount of information reduced. It was assumed that the greater the amount of mental processing or thinking required by the task the more it would show a decline in performance with speeding.

To test this hypothesis, a constant stimulus input consisting of eight numbers (48 bits) was presented, with the interstimulus interval varying from 4 to 1 sec. Six different information reducing transforms were studied varying from 0 to 40·3 bits reduction. These involved various combinations of relatively well-learned component operations including recording, addition and classification into such categories as high or low and odd or even. The results of several different experiments showed that the greater the information reduced between input and output the more performance declined with increasing speed. The relation between rate of decline in performance and information reduction was roughly linear.

Arithmetic operations

Two other investigations have been made applying information measures to predict the difficulty of arithmetic transformations. Thomas (1963) presents a constellation hypothesis of calculation. The general view taken by Thomas is that the time required for information processing is proportional to the amount of information which must be processed. This is in complete accord with the assumptions involved in the work discussed above. Thomas attempts to predict the difficulty of addition and multiplication by taking the logarithm of the sum of the values of each stimulus element and the response. For example, the equivalent information processed in adding $5+3 = 8$ is given by $\log(5+3+8)$. There is no *a priori* reason for this particular method which is, of course, not a calculation of information in the usual sense. Thomas finds that this calculation fits a variety of simple arithmetical problems which he has studied. Since the method uses the actual numerical value of the digits rather than the population from which they are drawn it provides differential prediction for many more situations than does the information reduction hypothesis. Thomas finds some subsidiary assumptions necessary in order to handle multiplication problems. Even with such additional assumptions the method is pretty well limited to basic arithmetical operations.

Wiegand (1963) also attempts to predict the difficulty of mathematical operations including addition and multiplication. He compares single-step and multiple-step operations. Wiegand's method considers separately the information in the stimulus, each successive step in the intervening operations and in the response. The basic

assumption is that each step obeys the relation between information transmitted and speed normally found in conservation tasks (Hick's law). He finds linear relations between the amount of information processed and the time to do the processing.

These two methods are quite similar to the information reduction notion in initial assumptions, though they differ in the method of quantification. This difference arises partly out of the fact that both of these methods are limited to tasks which, like serial arithmetic, allow specification of the steps or stages through which the subject must pass to solve the problem. Fitts (1959) has suggested that the amount of mental processing and, thus, performance in a serial task is a direct function of the number and complexity of such steps. However, in many tasks, it is difficult to specify the sequence of steps through which the subject must pass. For example, in learning a classification it is not, at present, possible to state each step which intervenes prior to the emergence of the correct response. The use of amount of information reduction allows the prediction of task difficulty solely as a function of input and output without the need to specify each step in the process. For this reason the measure is of greater generality and can be used to examine tasks which appear to be quite different. Presumably the utility of this method rests on the fact that it is sensitive to the amount of processing which intervenes between input and output. Thus, in mental arithmetic and other tasks for which it is presently possible to specify the intervening steps, additional methods of analysis such as those discussed in this section are valuable supplements to the use of transformation size in terms of amount of information reduction.

Classification learning

While the transformations required by simple arithmetical operations are convenient to study, they represent only a small fraction of the type of information transforms important in sequential tasks. The use of amount of information reduction as a predictor of task difficulty is particularly valuable because it has been extended to the use of classification rules (Posner, 1962, 1964*a*, *b*). Elsewhere I have reviewed a large number of studies of the learning of classification rules and their utilization after extensive learning (Posner, 1964*b*). The results of this review indicate that there are two basic types of information reducing transforms. In the first type, aspects of the stimulus input may be ignored in making the classification. This type represents selection of information. For example, if a subject is asked to classify playing cards by colour, the suit and value may be completely ignored. This class of task is said to involve *gating*. The second type requires every aspect of the input stimulus to be processed and represented in the response, but in a condensed form, such as in the arithmetic operations discussed previously or in the classification of two digit numbers into high and odd, low and even, etc. These are called *condensation* tasks.

During the learning of a classification rule the amount of information reduction is directly related to the length of time required to learn for both gating and condensation tasks. This generalization applies to multivariate concept learning (Archer, Bourne & Brown, 1955; Metzger, 1958) in which the concept is defined by a number of relevant dimensions. This is true both when the number of responses allowed match

the number of relevant stimulus levels so that the task only involves gating (Archer *et al.*, 1955; Battig & Bourne, 1961) and when the information is distributed across dimensions so that condensation is required (Shepard, Hovland & Jenkins, 1961). In most cases the relationship between amount of reduction and performance is linear.

The same general conclusion has also been found to apply using univariate stimuli in which the classification must be learned by rote (Metzger, 1958; Posner, 1964*a*). In one study (Posner, 1964*a*) the stimuli consisted of patterns of dots for which subjects learned to give pairs of related patterns the same name in a rote transfer design. The rate at which an original pattern could be related to each of its distortions in a concept assigned a single name was an inverse linear function of the amount of uncertainty between the two patterns. The uncertainty in this situation is calculated from the statistical rule which produces the distortions from the original pattern. Just as in the multivariate case the greater the amount of information encompassed by a common category the greater the time to learn the category. The reason for this correspondence between experiments in multivariate concept learning and this study of rote classification learning seems to lie in the close relation between perceived similarity and the amount of uncertainty between patterns. Posner (1964*a*) has shown that subjective similarity may be predicted with considerable accuracy from the uncertainty between an original pattern and its distortions. Thus similarity in univariate classification learning can be reduced to the same measure, amount of information, as has been used in the study of multivariate concept learning. When this is done the commonality between these experiments becomes apparent.

Utilization of concepts

Studies of the speed of classification of stimuli once the rule is already relatively well learned continue to show that the difficulty of condensation tasks is directly related to the amount of information reduced by the classification (Crossman, 1953; Pierce & Karlin, 1957; Pollack, 1963). The studies of arithmetic combination and classification already discussed show this relationship. However, gating tasks do not always increase in difficulty with the amount of information reduced once the rule governing the classification is well learned. For example, Archer (1954) showed that increasing the number of irrelevant dimensions does not vary significantly the speed of classification. In this situation the exact finding seems to depend on the degree of distinctiveness between the irrelevant and relevant information. If the material is well coded, gating information by use of a well learned rule does not seem to increase the task difficulty.

Summary. It is possible to predict the difficulty of many transformations from a knowledge of the informational properties of the task. A number of systems for doing this are available each having certain advantages and limitations. Perhaps the most general method uses as a measure the amount of information reduced in producing the required output. This measure does not require specification of the steps in arriving at the results. If this measure is used, the amount of information reduction is linearly related to task difficulty in a large number of tasks involving learning and utilization of classifications and simple arithmetic combinations. The ability to describe the difficulty of such transforms allows one to deal with the transformational

or thinking aspects of many complex sequential tasks. Before turning to those complex tasks, it will first be necessary to discuss the interrelations between such transforms and the storage or retention function.

III. Memory and thought

Few experiments, even among those normally called 'memory tasks', find the human subject operating in the pure retention mode. Rather the subject uses his past experience and his knowledge about the type of response he will have to make in order to reduce his storage load. Usually if the experimenter is interested primarily in the pure retention mode he does everything possible to avoid this. Thus he uses ordered recall which forces the subject to conserve the input information, uses difficult to organize strings of unrelated items, presents different material to several sensory channels simultaneously, or attempts to block the central processing capacity of the subject with other tasks. Ordered recall represents the pure retention end of a memorial dimension. This dimension reflects the degree to which stimulus information must be conserved in the response. As one moves away from ordered recall the subject is allowed to represent rather than reproduce the input in the output. Because of this it should be expected that such tasks would show increased emphasis upon transformations of various sorts.

Free recall and recognition

That this is the case is illustrated by the free recall task where the subject is not required to preserve order information. Garner (1962) has noted the close relationship between free recall and concept or classification learning. Virtually any relation between items of the list will lead to clusters or groupings, showing that the subject is making transformations of the list in accord with formal properties of the items. Even when the experimenter constructs lists of deliberately unrelated items, Tulving (1962) has shown that over successive trials the order of recall is not random but increasingly stereotyped, thus tending to reduce the information in recall lists below that in the random presentation. The Tulving result gives powerful evidence that the free recall situation always involves transformations of order by the subject in accord with formal similarities among items or with his past experience. Tulving's efforts also indicate that the extent of this organization, or in our language the size of such transformational processes, can be measured by successive increases in the amount of information reduced as the order becomes more stereotyped.

Just as free recall allows the subject to perform information reducing transformations through changes in order and still meet the task requirements, so recognition memory allows the subject to encode only critical parts of the stimulus input and still successfully perform his task. As would be expected, whenever the complexity of the input is sufficient to allow it, evidence exists that subjects encode only a part of the stimulus complex. For example, Anderson & Leonard (1958) have shown in studies of memory for complex forms that statistical rules which introduce symmetry serve to aid reproduction but have the opposite effect on recognition. They interpret this as meaning that such rules reduce the distinctiveness of features which can be selected for recognition. Dale & Baddeley (1962) have shown that increasing

the similarity of items on the recognition list to the one in store decreases performance. Presumably this is because the effectiveness of selection of relevant characteristics from the stimulus is reduced.

Interaction of transforms and retention

Both free recall and recognition studies illustrate that the degree of information conservation required by a memory task affects the form and character of the output. While those tasks which require information conservation are less susceptible to complex encoding strategies than those which, like free recall and recognition memory, allow the opportunity for such transforms, even in the memory span type of experimental situation transformation may play an important role in retention. Miller (1956) cites a study in which a single subject learned to recode binary numbers into octal digits and was, therefore, able to increase his retention from 12 to 40 digits. This demonstration requires a large amount of prior learning and, while Miller did not attempt to study this, must require a relatively slow rate of presentation since the subject must continuously perform the transform. That is, the process of transforming information from one coding system to another is a process which requires time. While pure retention shows a slight increase in efficiency as materials are presented at increasing rates (Conrad & Hille, 1958), at least within the range of easy perception, the transformation processes will show large decreases. Thus a transform which may be efficient in reducing the storage load is paid for by imposing a limitation upon the rate at which a subject can perform the overall task.

A recent study illustrates this trade off between the retention and transformation components of sequential tasks (Posner, 1964*c*). In this study subjects were presented with 8 digits at interstimulus intervals of either 2 or $\frac{1}{2}$ digits per sec. Two tasks were used; recall of the digits in the order presented or recall of the last four followed by the first four. It was found that this simple information conserving transformation of input order greatly improved performance when the rate of presentation was slow, but had no effect at the fast rate. Even this simple transform on order could improve performance provided the input information was sufficiently slow so that the transform itself did not interfere with the retention of items already in store.

What then are the effects of transformations of a more complex nature such as the combinations and classifications which might be imposed in problem solving or reading? In the material discussed so far emphasis has been placed on the fact that the memory load could be reduced in tasks which allow a subject to store information in coded form. However, the act of transformation may itself have systematic effects upon material in store when the transform occurs. To begin to uncover these relationships we undertook a series of studies (Posner & Rossman, 1964) of relatively simple tasks which involved a combination of retention and information reducing transformations. The results revealed some of the fascinating interrelations between these two processes at the heart of intellectual performance.

In the first study subjects were presented with eight random digits at a rate of 2 sec per digit. There were four groups of subjects each assigned to a given type of information transformation task. The tasks were as follows. (i) Reversal, in which the subject recorded each transformed pair in the reverse order. (ii) Addition, in which the subject added the digits of each transformed pair. (iii) Two bit classification,

in which the subject classified each pair into high-odd, low-odd, etc. (iv) One bit classification, in which high and odd or low and even comprised one category and the reverse the other category. These tasks were 0, 2·8, 4·6 and 5·6 bits reducing respectively. Within each transform group, subjects were required to transform the last pair only, last two pair and last three pair in various series. After the classification they attempted to recall all of the numbers in the order presented. Only the first pair of digits which was never itself transformed was scored. The results were scored in terms of the average increase in recall errors per transform. This score was 6, 14·8, 23·5 and 32 for the 0, 2·8, 4·6 and 5·6 reduction tasks respectively. Thus the amount of foregetting caused by a transformation increased with the size or difficulty of the transform. This relationship is shown in Fig. 1.

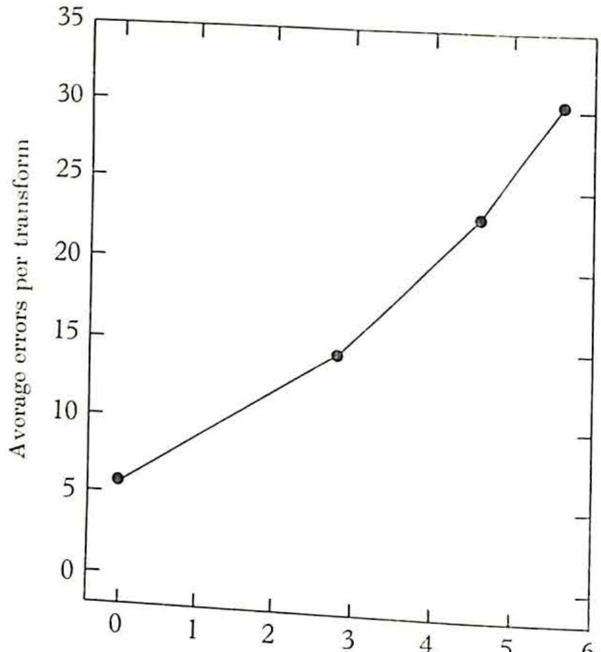


Fig. 1. Average increase in errors of recall per transformation as a function of the size of the transformation.

These effects are not due merely to the increased time in store which occurs with a difficult transform. In another study subjects were required to perform transforms of various sizes for a fixed period of time (10 sec) after a burst of material which was to be retained. Since the intervening task was self-paced subjects actually processed many more digits when the transform was simple. However, the percentage of errors in recall of the series to be retained increased fourfold from the simplest to the most difficult reduction task. In still another study the joint effect of time in store and transform difficulty was investigated. The result of this study was described by a family of diverging curves with errors increasing as time in store was changed from 0 to 30 sec and with the rate of increase governed by the difficulty of the transform. With interpolation of a classification task (4·6 bit reduction) the amount of forgetting after 5 sec was greater than for a recording task (0 bit reduction) after 30 sec

in store. Thus the effect of time in store depends on the difficulty of the task which is being interpolated.

These data fit rather well with a functional model of human processing in which the subject has a limited central capacity for processing information (Broadbent, 1958). Many studies have shown that retention in short-term memory depends upon a rehearsal process. Apparently this rehearsal process requires a portion of the central capacity, at least for digits, letters and similar materials. The portion of that capacity used in a transformation of incoming material is no longer available for rehearsal and retention of prior information. The extent to which an interpolated task requires such capacity is a function of its difficulty which can, at least for these tasks, be measured by the amount of information reduction required by the transform. To speak more simply the rate of forgetting depends not so much on what material occurs after an item but on how much processing the subject must do with that material.

The overall results of these experiments are important because they indicate the intimate relationship between thought and memory. While transformations of material occur in almost all memory tasks and are necessary because they reduce the amount of information that must be stored, the act of transformation causes a loss of material in short-term store when the transform occurs. The amount of the loss is related to the size or difficulty of the transform. It should be no surprise, therefore, that recent studies of concept learning (Cahill & Hovland, 1960; Hunt, 1962) and of serial reasoning problems (Simon & Kotovsky, 1963), have placed increasing emphasis on the importance of retention in the task.

Summary. So far we have reviewed evidence attempting to show the following: First, that a variety of transformations can be described in terms of their informational properties so that an *a priori* task difficulty can be specified. Secondly, that such transformations are related to closed system thinking (Bartlett, 1958) and thus size of transform can reasonably be thought to be representative of the amount of mental processing required by the task. Thirdly, that many so called memory tasks involve both retention and informational transformations (thought) working in close relation. Finally, that the performance of a transform has systematic effects upon material in store at the time of the transform.

IV. Application to complex tasks

The study of the pure processing modes of retention and transformation provides the basic materials which may allow analysis of very complex intellectual activities more akin to the adventurous and scientific thought about which Bartlett (1958) makes a number of interesting suggestions. In the balance of this paper an effort is made to analyse some basic questions within the general areas of induction, problem solving and reading in terms of information reducing transformations and immediate memory.

Induction

In discussing the boundary between closed and open system thought Bartlett (1958) suggests a type of problem in which 'at every decision we make, with every consequent move..., fresh information, or more evidence is made available, and

although descriptively this is just the same as it would have been if we had made a different decision, its significance relative to the objective is very closely dependent upon the moves and the extent of the moves, that we have already made' (p. 100).

The systematic and quantitative study of such a situation has recently had considerable emphasis in the area of decision making (Edwards, 1962). In this experimental situation subjects are required to select one of a number of alternative states of nature as being correct or most likely on the basis of information arriving sequentially from the environment. This paradigm is one which closely simulates that aspect of thinking which has been called induction. Since the subject must use each piece of evidence to provide information on the correct hypothesis, it clearly fits the description which Bartlett suggests above. This task still lies within the closed system, however, because it concerns selection among already formulated hypotheses rather than the actual formulation of the hypotheses themselves which would more closely fit the creation model.

This section will try to show that the induction task can be described as one in which a subject is required to make successive informational transformations as the result of incoming events. It will be shown that the size of such transformations as measured by the average amount of information reduced is inversely related to the efficiency of human performance in the task. Thus, induction can be viewed in terms of transformations on stimulus input and such a view not only demonstrates the commonality between this task and other intellectual performance but leads to predictions which are confirmed by experiment.

Semantic information. Bar-Hillel & Carnap (1953) suggest that the amount of uncertainty that one has about a set of hypotheses concerning the state of nature can be described in the same way as one describes uncertainty concerning which of 'n' events will occur. They chose to call this uncertainty 'semantic' since it refers not to the probability of occurrence of events but to the probability of the hypothesis being correct. Thus at any moment in time, if one can specify a set of hypotheses about the state of nature, each with a probability, it is possible to compute a quantity which refers to the total uncertainty or potential information at that time.

Statistical decision maker. The next problem in the description of induction is how a subject's uncertainty changes in the light of additional environmental information. Edwards (1962) and Watanabe (1960) have proposed the use of a statistical decision maker who transforms *a priori* probabilities by use of Bayes's Theorem. This statistical decision maker can be said to extract all the information possible from each environmental event so as to take maximum advantage in reducing uncertainty. It is proposed, not as a description of actual performance, but as a standard against which to compare human behaviour.

Serial induction. In a series of investigations Edwards and his co-workers (Edwards & Phillips, 1963; Phillips, Hays & Edwards, 1963; Edwards & Kramer, 1964) have compared the performance of human subjects with that of the statistical decision maker for problems in which subjects were required to induce the state of nature from a set of environmental events. In their simplest form the problems are as follows. Each subject is told that there are two bags which contain poker chips of two colours. The subject is told the proportion of chips of each colour in the two bags. For example, bag 1 may contain 0·6 red chips and 0·4 blue chips while bag 2 contains

0·4 red and 0·6 blue. The subject is also given an *a priori* probability that each bag is being used on a particular problem as, for example, the two bags may be equally probable at the start. The problem consists of a series of draws from one of the two bags with the subject rating the probability of each bag being in use after each draw.

The general conclusion from these investigations is that humans find it difficult to extract as much information from each event as could be extracted by use of Bayes's Theorem. This results in subjects consistently underrating the probability of the most likely bag. For example, in one sequence when the probability calculated by the statistical rule is 0·97, the subjects' ratings lie between 0·55 to 0·85. This finding of conservatism appeared in all the studies which required subjects to rate probability. In their most recent study, however, Edwards & Kramer (1964) have found a new and, from our point of view, basic phenomenon. In this situation subjects were allowed to buy information before reaching a decision which could bring varying amounts of payoff. In the two different sets of problems the events were drawn from a population in which their probabilities were either 0·7 to 0·3 or 0·6 to 0·4. When the performance of subjects in these two conditions is compared with that of a statistical decision maker obeying Bayes's Theorem and terminating at a point which maximizes expected payoff, it is found that in accordance with previous work on probability ratings subjects in the 0·7 to 0·3 condition required too much information before terminating. Those in the 0·6 to 0·4 condition, however, terminate more rapidly than the standard. Thus for the first time in Edwards's work there appears a condition in which the humans tend to reach a conclusion too quickly and, therefore, are not conservative. We will postpone the discussion of this finding until after a brief consideration of some other work with closely related results.

Table 1. *Illustrative induction problem: numbers directly under hypotheses are the* a priori *probabilities of the hypotheses; numbers under the events are the probabilities of the events given each of the hypotheses*

Hypotheses	1 (0·4)			2 (0·1)			3 (0·5)		
Events	A	B	C	A	B	C	A	B	C
	(0·3)	(0·3)	(0·4)	(0·2)	(0·7)	(0·1)	(0·6)	(0·3)	(0·1)

Prior to this last study by Edwards & Kramer (1964), Wallsten (1964) and I began to study the induction question. Our notion was to view induction like other information reducing tasks but as one in which the subject uses incoming information to reduce continually his uncertainty about the state of nature. Following Edwards, we used problems of the same type but with a somewhat wider range of conditions. In each problem the subject is shown a card which lists from two to four rules. Each rule is defined by the probability of occurrence of from two to four events. A typical problem is shown in Table 1. In each problem the subject continues to receive information in the form of events until he terminates, which he is instructed to do when he feels he has had sufficient information to reach a decision. The subject neither pays for information nor wins points for being correct, thus he must provide for himself a means of trading off number of trials for errors.

Our results have shown several interesting facts about the nature of human performance in this situation. First, in agreement with Edwards's work it is clear that humans can put together sequential evidence and reach a decision based upon such evidence. This is true even when there are as many as four rules each with four events to consider, and even when the subject is not able to look back on previous events in the series. Moreover, it looks possible from our results to tell what makes a particular inductive situation difficult and what produces subjects' conservatism.

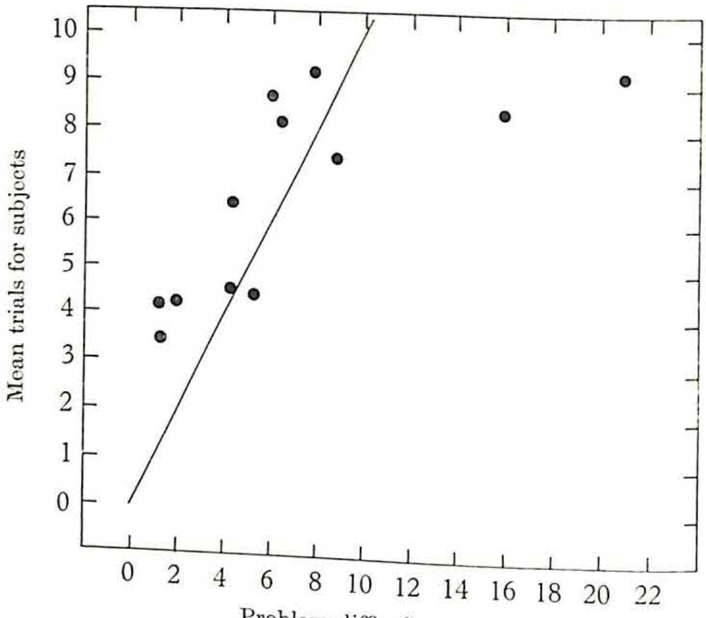


Fig. 2. Mean trials to termination for nine subjects as a function of the number of trials required by the standard statistical decision maker. The line represents perfect agreement between subject and standard. Points to the left of the line represent conservatism by subjects.

In Fig. 2 the general correspondence between a statistical decision maker and human performance is displayed. The abscissa represents the number of trials to termination of a statistical rule which makes transforms according to Bayes's Theorem and terminates when the rule, which is in fact correct, is first more probable than any other rule and never again becomes less probable. This termination criterion is the one which allows the standard to terminate most rapidly and still avoid error. Note that a close relation exists between the trials taken by the standard and the mean trials to termination of our nine subjects, indicating that the same factors which lead the statistical rule to require additional information also affect the general subjects' behaviour. That the rule which they induce is usually correct is seen from the fact that the error rates for subjects is only 18·5% at their places of termination.

It can be seen from Fig. 2 that as the number of trials that the statistical decision maker takes increases, the performance of subjects relative to the standard seems to be getting better. The full effects of this phenomenon are shown in Fig. 3. In this figure the mean percentage of uncertainty reduced is plotted against the difficulty of the problem for the statistical decision maker. As can be seen, subjects demand

virtually complete reduction of uncertainty before they terminate when the problem is easy, but much less when the problem is difficult. This says that when the average information reduction per transform is high for the statistical decision maker, subjects tend to be very conservative in their behaviour while, when the average size of transform is low, the subjects approach and even surpass the rapidity of the standard in termination.

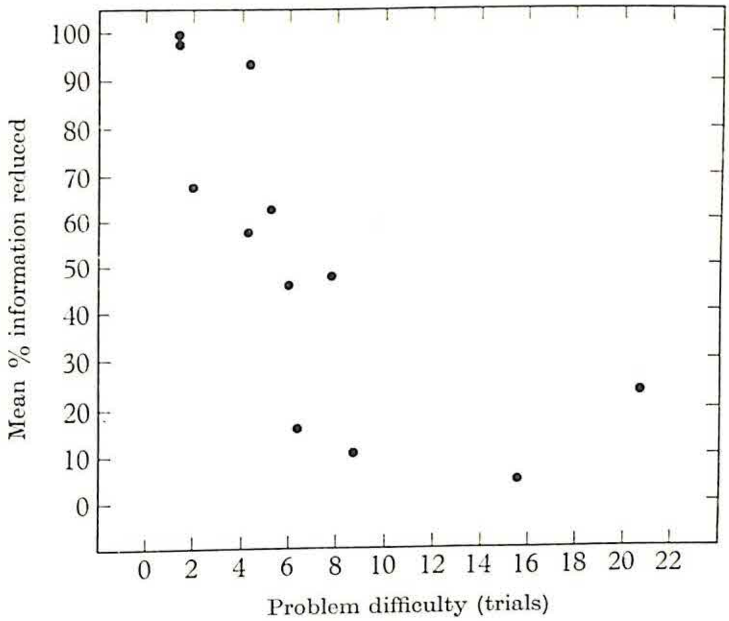


Fig. 3. Average percentage of information reduction prior to termination for nine subjects as a function of the number of trials required by the standard statistical decision maker to reach termination.

Sample size. Another type of induction experiment confirms the finding that the performance of subjects becomes increasingly inefficient as the size of the transform increases. These studies vary directly the size of a transformation by increasing the sample of information given on a single trial. When large amounts of information are supplied, the statistical decision maker can reduce uncertainty more than when a trial provides less information. In studies with quite different stimulus presentations the size of the sample available on a given trial was manipulated (Phillips *et al.* 1963; Peterson, Schneider & Miller, 1964) and it was found that subjects became increasingly conservative with respect to the standard as the transform became larger. This confirms the finding from studies which manipulate the average size of the transform within a sequential induction task.

Summary. These findings are only a beginning in the study of induction. However, they reveal the utility of viewing induction as an information reducing process which has much in common with other intellectual tasks. Like the finding in simple arithmetical transforms and in concept studies the greater the size of the information reducing transformations which a subject must perform the less efficient his performance becomes.

Problem solving

'Point of no return'. Bartlett (1958) suggests that one of the important aspects of thinking is its tendency to show a 'point of no return'. He contends that thinking, very much like other skilled activities, reaches a 'stage beyond which the further input of signals cannot produce a result because it fails to be noticed or does not produce a result because it is ignored. Alternatively, the new signals lead to a belated attempt to modify action and error follows' (pp. 17–18).

Bartlett presents a number of arguments and anecdotal demonstrations of this phenomenon. The bulk of his case rests upon a type of problem solving in which a subject is required to fill in the gap in a series of items related by a generation rule.

Simon & Kotovsky (1963) have investigated a number of such problems and attempted to specify the processing properties required of the human by means of a computer program. They find that the major variable contributing to task difficulty in such problems is the number of places in immediate memory which must be occupied simultaneously in order to derive the correct answer. They view the subject as sequentially encoding information from the series and storing this transformation. They argue that 'if a subject is able to extrapolate a sequence, he holds in memory something different from the bare sequence with which he was presented. The sequence, taken by itself, provides no basis for its own extrapolation.' He must then generate a pattern based on the transformed stimulus input. In this view such a problem represents exactly the combination of storage and information transforms with which this paper has been concerned.

Once a problem solving task is viewed as involving both retention and transformation the existence of a 'point of no return' follows as a natural consequence. This is so because the limitations in the amount of information which can be held in store forces the subject to select information from the stimulus relevant to the hypothesis upon which he is operating. Thus, he stores not the original series but a transformation of it which appears to preserve the information necessary for the solution. Of course, the more difficult the act of transformation the less material can be held in store. If the subject has the correct hypothesis he may not be hurt by elimination of information irrelevant to it. If, however, the hypothesis turns out to be wrong he no longer has the information available to modify his course of action. This is likely to lead to behaviour in which the subject sticks to a hypothesis even in the face of contradictory evidence, since he has no other real alternative.

Sequential presentation. If the raw information is no longer available when the subject modifies his hypothesis he is unable to obtain the information he needs to formulate a new position. This is the case in many studies of concept formation and doubtless accounts for the evidence that such studies are greatly affected by memory load (Cahill & Hovland, 1960; Hunt, 1962). This is so much so that models which consider the memory of previous information to be effectively zero after each trial can fit experimental data, particularly when the rate of presentation is high (Bower & Trabasso, 1963).

One problem solving situation which illustrates the effect on performance of manipulating the size of information transformations has been used by Whitfield (1951) in a pioneering attempt to apply informational analysis to problem solving.

Whitfield presented subjects with the problem of finding out how to classify each of eight objects into eight, four or two categories. Thus, in the terms being used in this paper, his conditions would involve information conservation, and information reduction of one and two bits respectively. Whitfield measured the difficulty of problem solution by the number of trials taken by subjects in excess of those which an ideal statistical observer would require to eliminate all but the correct classification. These values were 0·44, 1·2 and 3·8 trials per subject for the conservation, 1-bit, and 2-bit reduction conditions respectively. Whitfield further showed that one reason for the relative difficulty of the information reducing tasks is the fact that information derived in a given trial survives longest, i.e. is not violated on subsequent trials in the conservation task and shortest in the 2-bit reduction task. This would be predictable from the notion that retention is systematically related to subsequent processing.

Simultaneous presentation. Apparently even in those cases where the information remains available as the subject solves the problem the limitations of immediate memory may still be important. Thus, Simon & Kotovsky (1963) find that the difficulty of their serial problems is most closely related to the number of items which must be held in store simultaneously for a subject to infer the correct generating rule. This is true despite the fact that a written statement of the original problem remains in front of the subject. This means that in the analysis of problem solving, memory may be a limitation whenever a number of items must be integrated in order to produce the transformation which yields the solution. Having the items present on a page in front of the subject does not guarantee that the necessity for retention during the act of transformation is eliminated. Thus the analysis of problem solving as a sequential task involving retention and transformation may be extended to problems in which the subject has continual display of the raw information. At least this is a necessary conclusion from the results which Simon & Kotovsky (1963) present.

This view means that the 'point of no return' phenomenon can occur as a result of the interaction of retention and transformation even in those situations in which the display of raw information is continually present. Bartlett (1958) proposed that the effect of location of a gap in a series problem would be one appropriate test of the 'point of no return' notion in human thinking. Posner (1959) studied a number of series reasoning problems with static visual displays in order to test Bartlett's suggestion. In these problems the subject must fill the gaps by means of a rule derived from the portion of the series provided. Using forty-eight different problems involving numerical, alphabetic, pictorial and semantic materials it was found that there was significantly better performance with late gaps than with early gaps, though the problems were otherwise identical. For example, the problem 6, 9, 5, 8, 4, 7——1, 4, 0 (solution 3, 6, 2, 5) was solved by about twice as many subjects (23 *vs.* 12) as the problem 0, 4——3, 7, 4, 8, 5, 9, 6 (solution 1, 5, 2, 6). This phenomenon requires further study, but may be related to the number of items the subject takes into account in solving the problem.

The results of this section indicate the utility of a view of problem solving in terms of the general laws of retention and transformation. By use of this approach it may be possible to realize Whitfield's (1951) goal of 'providing an objective form of difficulty, associated purely with the problem and not dependent on behaviour'.

Reading

Perhaps the most important serial task which human beings perform in terms of its frequency in daily life is the extraction of information from written and spoken material. It is clear that neither reading nor listening can be thought of as a pure retention task. When someone asserts that he remembers what is in a book or story he does not mean that he can reproduce it in ordered recall. Rather he means that he can state the main ideas, relationships, or concepts asserted by the story.

In coming to understand materials of this sort a subject must impose upon the words he hears or sees transformations which include selection, combination and classification of the sort we have discussed earlier. To the extent that such transformations are involved they should show up as in other memory tasks by reducing the rate at which a subject can assimilate information.

Rate of reading. Poulton (1958) investigated the effect of increasing the rate of presentation in reading. He was interested in comparing recall of specific words (pure retention) with a task where the subject had to select which of two statements had the same meaning as the one he had previously read. The first task approximates information conservation since it requires virtually complete recall, though Poulton did not require preservation of order information. The second task clearly allows more opportunity for the kind of information reducing transformations which have been reviewed in this paper. The subjects were required to read at rates of 293, 146, 73, and 37 words per minute. What is important for this discussion is that increasing the rate of reading resulted in significantly less material being understood as measured by the ability of subjects to select statements with the same meaning, but did not have as great an effect upon the recall of words. Poulton concluded: 'these results suggest that there is a limit to the amount of material which can be understood in a given time. If we proceed faster than this we may be able to recall slightly more words, but we shall not be able to recall more meaning. To use the terminology of information theory, the rate of coding the information in reading appears to have an upper limit, which is reached before the rate at which information can be stored in an uncoded form.'

It appears possible to analyse reading into a storage and an information transformation component. Just as in results with simple memory tasks, the rate of presentation has less effect upon the storage function than upon the transformations. Unfortunately, it was not possible in Poulton's experiment to specify in any detail the characteristics of the information reducing transformations required to produce the material needed for answering the content questions.

Types of transforms. Dawes (1964) has attempted to develop methods by which the transformations in connected materials can be evaluated. It is not yet clear whether such transformations will lend themselves to quantification by means of information measures. Dawes constructed stories which established relations between various sets. He made an analysis of the stories in terms of Venn diagrams which represented the verbal relations of the story. Two basic types of relations between pairs of sets were possible. In nested relations one set was included in either the other set or in its complement, while in a disjunctive relation both sets have common and uncommon elements. The disjunctive relation requires the subject to

carry in store more categories than does the nested relation. Dawes postulated, therefore, that in retaining the story subjects would tend to produce more overgeneralizations (retain disjunctive relations as nested) than pseudodiscriminations (retain nested relations as disjunctive). This finding was borne out in the experimental studies.

The potential importance of this method is very great. Many sentences, particularly declarative sentences, can be said to assert relations between sets. The comprehension of a sentence can thus be viewed as a kind of miniature concept utilization experiment. For example, suppose it is asserted that 'all scientists support the test ban, but no one else does'. This sentence asserts an identity relation between the sets 'scientists' and 'test ban supporters'. When a subject is then required to evaluate a statement such as 'some non-scientists support the test ban', he is being placed in a kind of transfer situation. His comprehension of the sentence may be said to be greatest when he can utilize the concept presented in the sentence in the widest number of transfer situations. One then can begin to make predictions about the difficulty of assimilation of information from a sentence as a function of the number of sets and the complexity of their relations. Dawes's studies with stories are a beginning of the effort to analyse connected materials using some of the same principles and methods which have grown up in the study of concept learning with patterns and multivariate dimensions. In this way the difficult problem of comprehension of sentences can be examined in terms of the types of transformations required to utilize the concepts they assert in a set of defined transfer situations. At the present time we are conducting experiments with a view of formalizing this kind of analysis.

V. Conclusion

It is difficult to draw firm conclusions from work which is really in its very early stages. The major point of this review is that a number of aspects of human intellectual performance, which have resisted systematic investigation for so long, can be examined analytically in terms of the transformation processes and retention involved. Each of these pure processing modes can be studied in abstraction from the complexities of real life tasks. This approach to information handling is roughly analogous to viewing complex learning in terms of combinations of simpler S–R associations. Through experimentation it is hoped that the basic laws governing human capacities for transformation and retention may be found. Such experiments are bound to be artificial because they abstract the crucial processes of natural tasks rather than reproduce the tasks. It is, therefore, continually necessary to make certain that the processes which are being abstracted are ones which really play an important role in intellectual performance. Thus the direct study of processing in relatively complex situations is also necessary.

The data reported here seem to show that an analysis of intellectual performance in terms of informational transformations, immediate memory and the interactions of these processes can provide tools adequate to the analysis of significant questions in human thinking.

Much of the research reported in this article was supported by a grant from the Research Committee of the University of Wisconsin with funds provided by the Wisconsin Alumni Research Foundation. The article was written at the suggestion of Sir Frederic Bartlett, who was

kind enough to indicate that my research on informational approaches to thought might be of interest to readers of this journal. This paper was prepared during the summer of 1964 while the author was visiting research psychologist at the Human Performance Center, Department of Psychology, University of Michigan, supported by the Advanced Research Projects Agency of the Department of Defense and monitored by the Air Force Office of Scientific Research, under Contract No. AF 49 (638)–1235. It owes much to the assistance of Professor Paul M. Fitts.

References

Anderson, N. S. & Leonard, J. A. (1958). The recognition, naming and reconstruction of visual figures as a function of contour redundancy. *J. exp. Psychol.* **56**, 262–70.

Archer, E. J. (1954). Identification of visual patterns as a function of information load. *J. exp. Psychol.* **48**, 313–17.

Archer, E. J., Bourne, L. E. & Brown, F. G. (1955). Concept formation as a function of irrelevant information and instruction. *J. exp. Psychol.* **49**, 153–64.

Bartlett, F. C. (1958). *Thinking: An Experimental and Social Study*. New York: Basic Books.

Bar-Hillel, Y. & Carnap, R. (1953). Semantic information. *Brit. J. Phil. Sci.* **4**, 147–57.

Battig, W. F. & Bourne, L. E. (1961). Concept identification as a function of intra- and inter-dimensional variability. *J. exp. Psychol.* **61**, 329–33.

Bower, G. & Trabasso, T. (1963). Reversals prior to solution in concept identification. *J. exp. Psychol.* **66**, 409–18.

Broadbent, D. E. (1958). *Perception and Communication*. New York: Pergamon.

Cahill, H. E. & Hovland, C. I. (1960). The role of memory in the acquisition of concepts. *J. exp. Psychol.* **59**, 137–54.

Conrad, R. & Hille, B. A. (1958). The decay theory of immediate memory and paced recall. *Canad. J. Psychol.* **12**, 1–16.

Crossman, E. R. F. W. (1953). Entropy and choice time: the effect of frequency imbalance on choice response. *Quart. J. exp. Psychol.* **5**, 41–51.

Dale, H. C. A. & Baddeley, A. D. (1962). Alternatives in testing recognition memory. *Nature, Lond.*, **196**, 93–7.

Dawes, R. M. (1964). Cognitive distortion. *Psychol. Reps. Monogr. Suppl.* **14**, no. 4.

Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors*, **4**, 59–73.

Edwards, W. & Kramer, E. (1964). Actual and optimal information seeking strategies. Paper delivered to Midwest Psychological Association.

Edwards, W. & Phillips, L. D. (1963). Conservatism in probabilistic inference. Paper delivered to Midwest Psychological Association.

Fitts, P. M. (1959). Human information handling in speeded tasks. *Research Report RC* 109, *IBM*.

Garner, W. R. (1962). *Uncertainty and Structure as Psychological Concepts*. New York: Wiley.

Hunt, E. B. (1962). *Concept Learning*. New York: Wiley.

Hunt, E. B. (1963). Memory effects in concept learning. *J. exp. Psychol.* **62**, 598–604.

Metzger, R. (1958). A comparison between rote learning and concept formation. *J. exp. Psychol.* **56**, 226–31.

Miller, G. A. (1956). The magical number seven; plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81–97.

Morin, R. E. & Forrin, B. (1963). Response equivocation and reaction time. *J. exp. Psychol.* **66**, 30–6.

Peterson, C. M., Schneider, R. J. & Miller, A. J. (1964). Sample size and the revision of subjective probabilities. *Behaviour Research Lab. University of Colorado, Report*, no. 41.

Phillips, L. D., Hays, W. & Edwards, W. (1963). Conservatism in probability estimation. Unpublished manuscript, University of Michigan, Institute of Science and Technology.

Pierce, J. R. & Karlin, J. E. (1957). Reading rates and the information rate of a human channel. *Bell Sys. tech. J.* **36**, 497–516.

Pollack, I. (1963). Speed of classification of words into super-ordinate categories. *J. verb. Learn. verb. Behav.* **2**, 159–65.

Posner, M. I. (1959). The effect of terminal points and related factors in the solution of serial reasoning problems. Unpublished M.S. Thesis, University of Washington.

Posner, M. I. (1962). An informational approach to thinking. *Air Force Office of Scientific Research Report*, no. 2635, ASTIA Document 276136.

Posner, M. I. (1963). Immediate memory in sequential tasks. *Psychol. Bull.* **60**, 333–49.

Posner, M. I. (1964*a*). Uncertainty as a predictor of similarity in the study of generalization. *J. exp. Psychol.* **68**, 113–8.

Posner, M. I. (1964*b*). Information reduction in the analysis of sequential tasks. *Psychol. Rev.* **71**, 491–504.

Posner, M. I. (1964*c*). Rate of presentation and order of recall in immediate memory. *Brit. J. Psychol.* **55**, 303–6.

Posner, M. I. & Rossman, E. (1964). The effect of size of interpolated information reducing transforms upon immediate recall. *J exp. Psychol.* (in the Press).

Poulton, E. C. (1958). Time for reading and memory. *Brit. J. Psychol.* **49**, 230–45.

Shepard, R. N. (1963). Production of constrained associates and the informational uncertainty of the constraint. *Amer. J. Psychol.* **76**, 218–28.

Shepard, R. N., Hovland, C. I. & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychol. Monogr.* **75**, (whole no. 517).

Simon, H. & Kotovsky, K. K. (1963). Human acquisition of concepts for sequential patterns. *Psychol. Rev.* **70**, 534–46.

Thomas, H. B. G. (1963). Communication theory and the constellation hypothesis of calculation. *Quart. J. exp. Psychol.* **15**, 173–91.

Tulving, E. (1962). Subjective organization in free recall of 'unrelated' words. *Psychol. Rev.* **69**, 344–54.

Wallsten, T. S. (1964). Human performance in inductive information reduction. Unpublished Honors Thesis, University of Wisconsin.

Watanabe, S. (1960). Information theoretical aspects of inductive and deductive inference. *IBM J. Research and Develop.* **4**, 208–13.

Whitfield, J. W. (1951). An experiment in problem solving. *Quart. J. exp. Psychol.* **3**, 184–97.

Wiegand, K. L. (1963). Information theory and human behavior: Uncertainty as a fundamental variable in information processing tasks. *Tech. Doc. Report*, no. AMRL 63–89.

(*Manuscript received* 1 *September* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 217–222
Printed in Great Britain

RESPONSE PATTERNS AND STRATEGIES IN THE DYNAMICS OF CONCEPT ATTAINMENT BEHAVIOUR

By RIVKA R. EIFERMANN
Psychology Department, The Hebrew University of Jerusalem

Following Bruner, Goodnow & Austin an analysis of subjects' response patterns in conjunctive concept attainment was undertaken. The analysis was based, in addition to the card choices used in previous experiments, on subjects' verbal reports about each of their choices. The aim was to trace the dynamics of concept attainment behaviour and determine how systematic and consistent response patterns become over a series of problems. It was found, first, that there exists a positive correlation between the ability to justify one's card-selection verbally and efficiently in concept attainment. Secondly, a majority of subjects became systematic in their response pattern after being exposed to a few problems of the same type. Thirdly, a good number of subjects, after having attained systematic response patterns became unsystematic again for a shorter or longer run. The theoretical concepts of *holistic* and *elementaristic* strategies are disdinguished from the observational ones of *component-centred* and *concept-centred* strategies.

Bruner, Goodnow & Austin (1956) describe four 'ideal strategies' of behaviour in certain concept attainment situations, saying that these represent 'the way we would set the computer to do what the subject appears to be doing'. One of the drawbacks of their approach is that it is not always possible to differentiate between these strategies on the basis of the observed responses. In order to overcome this drawback, an experiment was devised in which the subjects were requested not only to *do* something (to select certain cards), but also to *report* on the reasons for their choices. It is commonly claimed that subjects who successfully solve problems which require logical thinking are often unable to explain their behaviour. Our assumption was that ability to express oneself adequately while performing a structured task which requires logical thinking would correlate positively with efficiency in performance, i.e. with performance which is most favourable in terms of the amount of information attainable at each step in the process of solution. An additional aim was to investigate the dynamics of concept attainment behaviour and in particular, to see if there would be an increase in how systematic the response patterns were over a sequence of similar problems.

METHOD

Material

Two sets of sixteen cards each, a Form and a Number set, were used. Each of the cards in either set consisted of a unique combination of four attributes, each represented by one of two values. Table 1 presents the attributes and their values.

Procedure

Twenty-five first-year psychology students were tested. They had no previous familiarity with the experiment. Ten of the subjects were assigned to the Form group (two men and eight women) and fifteen (six men and nine women) to the Number group. Each subject was tested individually, and had to discover four concepts of two attribute-values and four concepts of three attribute-values. These concepts were the same for all subjects and were presented in random order to each subject.

The task was explained to the subject in detail and it was emphasized that he should attempt to attain the concept with a minimum number of card-choices, and that 'all concepts would

consist of two or three attribute-values'. The subject was also requested to give his reasons after each choice, and was urged not to guess the concept. He was then told whether the card chosen did, or did not, instantiate the concept. The search for each concept started after all cards had been laid out in an ordered array in front of the subject and one card representing 'a positive instance' of the concept had been pointed out to him. The search was terminated as soon as the subject announced the correct concept.

Table 1. *Attributes and their values represented in unique combinations on sixteen cards: (a) the Form set, (b) the Number set*

Attributes							
(a) Form set				(b) Number set			
Shape	Colour	Size	Background	Units	Tens	Hundreds	Thousands
Square	Red	Large	Clear	---1	--1-	-1--	1---
Circle	Green	Small	Striped	---2	--2-	-2--	2---

RESULTS

Categories of response

Somewhat in line with the categorization of Bruner *et al.* (1956) the hypothesis was investigated that people may be divided with regard to their approach to concept attainment problems into what are here called 'elementaristic' and 'holistic' solvers. An 'elementaristic solver' concentrates his attention on the components exemplified by the first positive instance and aims to find out which of those components is, or is not, part of the concept. He may do so by testing one component at a time or by pairs, triplets, etc. On the other hand, a 'holistic solver' envisages a concept which he tests as a whole by selecting cards which instantiate that concept.

In order to test this hypothesis, 'responses', defined as pairs consisting of a *choice* and the accompanying *verbal report*, were categorized as follows ('differences from the first positive instance' in *choice* and 'changes' in *report*, always refer to attribute-values that either were not yet explicitly tested or else were explicitly retested):

1. *Choice* differs in just one attribute-value from the first positive instance.
   Report: comments on attribute whose value was *changed*.
   Examples of report: 'To see whether the 2 is important'; 'to eliminate the units'; 'now I "isolate" size' (with size being the attribute whose value was changed).
2. *Choice* differs in two attribute-values from the first positive instance.
   Report: comments on attributes whose value was *changed*.
   Example of report: there were no actual examples in the particular experiment described in this paper but in another, similar experiment: 'I am changing colour and shape at the same time, although I know it's a gamble'.
3. *Choice* differs in just one attribute-value from the first positive instance.
   Report: comments on all attributes whose values were left *unchanged*.
   Examples of report: 'To know whether the concept is 22–1' (22–1 were left unchanged); 'I am working on the hypothesis that the concept is 11–1'; 'I thought the concept might be 21–1'.
4. *Choice* differs in two attribute-values from the first positive instance.
   Report: comments on attributes whose value was left *unchanged*.
   Examples of report: similar to those presented in category (3), but comments on the two attributes whose values were left unchanged.

5. *Choice* differs in just one attribute-value from the first positive instance.
   Report: comments on an attribute whose value was left *unchanged*.
   Example of report: 'I want to "isolate" colour' (while the attribute whose value was changed was shape).
6. *Choice* differs in just one attribute-value from the first positive instance.
   Report: comments on *two* attributes whose values were left *unchanged*.
   Example of report: 'I am testing red and small' (while the attribute whose value was changed was background).
7. *Choice* differs in one or more attribute values from the first positive instance.
   Report: comments both on attributes whose values were *changed* and on attributes whose values were left *unchanged*.
   Example of report: 'I want to test "red square"' (while the attribute whose value was changed was colour).
8. *Choice*: any card.
   Report: contains no comments on any attributes whose values were either changed or left unchanged.
   Examples of report: 'I simply guess'; 'I don't know why I did it'.

In analysing the results, responses of types (1) and (2) taken together were distinguished from types (3) and (4) taken together on the basis of the differences in the reports in these two classes of response: (1) and (2) are 'component-centred' responses while (3) and (4) are '(whole) concept-centred' responses. In all these cases there is a one-to-one correspondence between choice and report, while no such correspondence obtains in the last four cases which will be called 'opaque'. Responses of type (8) were separated from the other opaque responses, since this was the only type of response which contained no specific comment on any attribute whatsoever.

Subjects sometimes 'guessed' the concept, contrary to instructions; i.e. they either named a wrong concept while already possessing information sufficient for identifying the correct concept (a 'redundant guess'), or else they named a concept before they had information sufficient for such an identification (a 'guess'). In this latter case, if what they said was wrong the experiment was continued; if correct, the experiment was terminated, but their solution was still counted as a 'guess'. Sometimes a subject made a choice to which the experimenter's reaction that it did or did not instantiate the concept was already determined by the subject's previous choices; such choices could not provide the subject with further information and were termed 'redundant'.

No difference in response patterns was associated with type of material and the results for Form and Number cards are therefore discussed together. Table 2 gives the mean numbers of responses of the different types for all twenty-five subjects, with responses to the first problem distinguished from responses to the other seven. The table shows that there were far more component-centred responses than concept-centred responses. In fact, all component-centred responses were of type (1), i.e. *single*-component-centred. The proportion of redundant responses was by far the greatest in the case of opaque responses.

Development of systematic response patterns in the group as a whole

In order to test whether response patterns become more systematic through successive problems, the number of deviations from the more frequent of the two major response categories (component-centred or concept-centred) was determined for each problem. Table 3 gives the number of subjects who deviated from their own

more frequent major response category, as well as the number of subjects who 'guessed' the concept once or more than once, for each of the eight successive problems. The number of subjects deviating from their own more frequent response category was by far the greatest on the first problem (sixteen out of twenty-five subjects). There was no decline, however, in the number of subjects guessing on successive problems.

Table 2. *Mean number of responses, mean number of redundant responses and proportion of redundant responses, per subject per problem, in four response categories for the first problem and for the second to eighth problems* ($n = 25$)

Problems		Responses: Component-centred	Concept-centred	Opaque: Types 5–7	Type 8	All	Guesses
First	Mean number of responses	1·72	1·36	0·48	1·12	1·60	0·40
	Redundant responses: Mean number	0·24	0·20	0·20	0·56	0·76	0·08
	Redundant responses: Proportion	0·14	0·15	0·42	0·50	0·48	0·20
Second to eighth inclusive	Mean number of responses	2·64	0·95	0·02	0·45	0·47	0·29
	Redundant responses: Mean number	0·15	0·09	0·00	0·14	0·14	0·09
	Redundant responses: Proportion	0·06	0·09	0·00	0·31	0·29	0·31

Development of systematic response patterns in individual subjects

Twenty out of twenty-five subjects reached a response pattern so systematic that all but at most one response belonged to the same response category. Of these, the responses of sixteen subjects were component-centred and all of type (1), i.e. single-component-centred, while the responses of four subjects were concept-centred. However, two of the subjects whose responses were component-centred responded unsystematically on a later problem (one on the fifth and the other on the sixth problem), but returned still later to the same systematic behaviour as before. Of the subjects whose responses were concept-centred one turned unsystematic on the third problem and another on the seventh; neither regained the degree of systematic behaviour which was apparent in their earlier performance.

Table 3. *Number of subjects who (a) deviated from their more frequent response category; (b) guessed on each of eight problems* ($n = 25$)

Successive problems	Deviations: No deviations	One deviation	More than one deviation	Guesses: One guess	More than one guess
1st	9	6	10	6	2
2nd	16	3	6	3	2
3rd	18	3	4	7	0
4th	19	4	2	5	0
5th	16	5	4	7	2
6th	13	7	5	3	1
7th	17	5	3	5	0
8th	19	4	2	7	1

Discussion

Categories of response

Results show that in the cases where responses were opaque, either of types (5, 6 and 7) or of type (8), redundant responses were also by far the greatest in number. Since this finding was not necessitated by logic, it suggests, that in our experiments at least, lack of ability to justify one's problem solving behaviour tends to go together with relative inefficiency in the attainment of information. It is not surprising, therefore, that the largest proportion of opaque responses occurred on the first problem presented.

This finding would seem particularly pertinent, since verbal report has been re-introduced of late into the investigation of thought processes (Newell & Simon, 1961; Wason, 1960). Wickelgren & Cohen (1962), in a situation comparable to the present one, have even taught their subjects a limited, artificial language for the purpose. In this way they have indeed by-passed some of the problems of opaque responses, but from our viewpoint they have at the same time given up information which is psychologically relevant.

The dynamics of concept attainment behaviour

It is not unreasonable to see in a consistent sequence of systematic response patterns a manifestation of a specific concept attainment strategy. Thus, the behaviour of those subjects who reached a systematic response pattern of the component-centred category and consistently stuck to it is to be regarded as good evidence for their acting in accordance with an 'elementaristic' strategy. Similarly, the behaviour of those subjects who reached a systematic response pattern of the concept-centred category, is to be regarded as indicating action according to a 'holistic' strategy.

Our classification corresponds approximately to the 'strategies' of Bruner *et al.* (1956), though these authors did not draw the distinction between the behavioural and the theoretical concepts. The elementaristic strategy is related to 'conservative focusing' (type 1) and to 'focus gambling' (type 2) and the holistic strategy approximates 'successive scanning'. However, since Bruner *et al.* did not have access to systematic verbal reports, it is difficult to see how they succeeded in their empirical classification of many choice sequences (see Eifermann, 1965, for examples).

Bruner *et al.*, who view strategies functionally, asking 'what the behaviour sequences accomplish for the organism', claim that 'shifts in strategies can be described and related to changes in the requirements of the tasks set'. Evidently, it was not functional to turn unsystematic and yet such changes did occur. Nor would it have been functional, from the viewpoint of task requirements, to shift from the elementaristic strategy which secures the greatest amount of information at each step to the holistic strategy. No such shifts occurred in this experiment, but they did occur in another similar experiment.

Thus, it appears that the development of a subject's concept attainment behaviour will not necessarily proceed along a smooth path towards functionally improved performance. It may be said that on the whole subjects tend to reach consistency and functional efficiency in performance. But the exceptions in their behaviours are not

few and they vary from the opaque responses—which may be due to factors such as inability, or indifference, or frustration—to shifts in strategy and abandonment of behaviour according to a strategy, due, perhaps, to a 'need for a change' or 'a desire to explore' or any number of factors far removed from immediate task requirements. The resourcefulness—or capriciousness—which is expressed in human behaviour even when dealing with the most 'dull' and 'abstract' kind of cognitive task, stands out too clearly to be ignored.

Hence, if we are to construct a comprehensive model of human performance, even for situations which require 'pure' logical thinking only, it is essential (*a*) to reach and (*b*) to analyse such 'non-logical' factors and treat them as relevant aspects of our data. This was done, for example, by Donaldson (1960), and Gyr, Thatcher & Allen (1963). A few steps in this direction of exploration and analysis were also undertaken in the present experiment in the treatment of opaque responses, and in the analysis of dynamic aspects of concept attainment behaviour. These aspects can now be examined in greater detail, and the following questions, among others, can be put to experimental test: What kinds of factors affect consistency in performance even though task requirements remain unaltered? To what extent can motivation affect consistency? How far can people who have at a certain point reached a consistent sequence of systematic response patterns be expected to retain such behaviour afterwards? To which other kinds of problems, if any, would this systematic concept attainment behaviour transfer? So long as we do not have answers to questions like these, the relations between such theoretical concepts as strategy and such observational concepts as systematic response patterns remain underdetermined.

Grateful acknowledgements are due to Mr Y. Giveon for the important part he took in the conception of the experiment, and to Miss Daphna Agmon who participated as experimenter. Mrs Ruth Khleif-Steinitz's assistance in the analysis of the data was invaluable. I am also grateful for her enlightening comments on the manuscript, as for those of Dr C. Greenbaum and Mr O. Schild of the Psychology Department, and of Prof. Y. Bar-Hillel of the Philosophy Department, the Hebrew University of Jerusalem, and of Dr P. C. Wason of the Medical Research Council Industrial Research Unit, University College, London.

REFERENCES

BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

DONALDSON, M. (1959). Positive and negative information in matching problems. *Brit. J. Psychol.* **50**, 253–62.

EIFERMANN, R. R. (1965). Selection strategies in concept attainment: a re-examination. In *Scripta Hierosolymitana: Studies in Psychology*. (ed. R. R. Eifermann). The Hebrew University of Jerusalem: Magnes Press.

GYR, J., THATCHER, J. & ALLEN, G. (1962). Computer simulation of a model of cognitive organization. *Behav. Sci.* **7**, 111–16.

NEWELL, A. & SIMON, H. A. (1961). Computer simulation and human thinking. *Science*, **134**, 3495.

WASON, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quart. J. exp. Psychol.* **12**, 129–40.

WICKELGREN, W. A. & COHEN, D. H. (1962). An artificial language and memory approach to concept attainment. *Psychol. Rep.* **10**, 815–27.

(*Manuscript received* 20 *January* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 223–232
Printed in Great Britain

THE PLACE OF SUBJECTIVE EXPERIENCE IN CONTEMPORARY PSYCHOLOGY

By ALAN RICHARDSON
University of Western Australia

This paper discusses some problems raised by Koch (1958) concerning the role of experiential analysis in contemporary psychology. More specifically the attempt is made to distinguish the concept of experience from the concept of behaviour and to justify the use of subjective report on both logical and empirical grounds. The problem of developing more accurate categories for self-observation and of developing valid constructs from self reports is considered. The final section of the paper is concerned with the theoretical implications of the experientially derived construct of imaging ability and of the subjective activity of mental practice.

The point of departure for this paper is the brief but important commentary on experiential analysis by Koch (1958) in his epilogue to the first three volumes of Project A.

The following quotations* provide an appropriate frame of reference for the specific questions to be discussed: '...for more than forty years behaviourist epistemology has had the *pragmatic effect* of fostering a set of attitudes which tend to either devalue or divert attention from most problems which, by virtue of historical or extra-scientific associations, have an "experimental odour"—quite independently of whether the investigator believes the problem compatible in principle with behavioristic methods'. However: 'An important and quite general trend of the essays is an increased recognition of the role of direct experiential analysis in psychological science' (p. 766). Nevertheless: 'there has been no marked tendency among authors in the present study...to join in any explicit way the many methodological and empirical questions that might be asked concerning fruitful and rigorous utilization of experiential data. Certainly the manifold current stresses against behaviourist epistemology invite such questioning. Yet issues concerning optimal techniques for experiential observation, the formulation of adequate dependent variable categories, the integration of behavioural and experiential data, the construction of theoretical concepts from experiential data, etc., have been addressed by indirection, if at all' (p. 768).

It is with these general comments in mind that the following specific questions have been raised. These questions have not been dealt with exhaustively nor do they exhaust the important questions to be asked in this area. The topic is of sufficient importance to require that a start be made.

(1) Can a valid conceptual distinction be made between the class of psychological events labelled behaviour and another class labelled experience?

(2) What is the logical and empirical justification for the utilization of subjective report?

* From *Psychology: A Study of a Science*, Vol. III, ed. S. Koch (copyright 1959), McGraw-Hill Book Company and used with their permission.

(3) How can reliable and valid independent and dependent variables be constructed from the raw data of the subjective report?

(4) Is there evidence that any experiential variables might have widespread theoretical significance?

BEHAVIOUR AND EXPERIENCE

Despite the powerful influence of classical behaviourism and its offshoots it was not until some time in the late 1930's that the concept of experience seems to have been defined out of existence by the authors of most general texts in psychology. Until that time psychology was usually defined as the scientific study of behaviour and experience but it is rare to find a general text written in English, after 1940, that retains the concept of experience in its definition.

The main argument that was said to justify its removal may be summarized as follows: if by the term experience is meant some conscious event that is essentially private and incommunicable then such an event has no place in a science of psychology. If, on the other hand, the term experience refers to a class of conscious events that can be communicated in a written, spoken or key-pressing form, then it has become just like all other observable behaviour so that the continued use of the term 'experience' is simply redundant.

The fallacy in this argument results from the term 'behaviour' having two distinct meanings. First, as the antithesis of experience, it refers to all forms of overt action that are of interest in their own right (i.e. conceptual behaviourism). Secondly, it refers to a tangible, observable set of dial readings, pen-recorder tracings, check marks on questionnaires, tape recordings or whatever, that serve as indices of some process or state which may be either behavioural or experiential (i.e. methodological behaviourism).

This distinction is important. When a subject makes overt movements with a pencil to produce ratings on a questionnaire of imagery modes, it does not mean that our research interest is focused on these overt movements. It is not the motor behaviour that is of interest but the subjective experience of having imagery. It just so happens that one method of measuring preferred mode of imagery involves such motor behaviour for its communication.

Subjective report is not the only method by which indices of experiential states or processes may be obtained, though initially it will always be necessary to validate these 'objective' measures by reference to subjective reports at some stage in the research. If at some later stage it becomes possible to locate a particular experiential construct within a nomological network then further dependency on subjective report for that construct may no longer be necessary. In this regard it is of interest to note Cronbach & Meehl's (1955) reminder that 'Teacher judgements once constituted the criterion against which the individual intelligence test was validated' though today, with the aid of factor analysis, we have succeeded in lifting ourselves by our bootstraps. In summary, we must not confuse the utilization of behaviour as a method with an exclusive interest in behaviour as a class of psychological events to be described and explained.

Having stressed the importance of making a conceptual distinction between behaviour and experience it will be appropriate to give some examples of what is

meant by an experiential problem. Such problems are those that might be summarized under the heading of self-conscious states or processes of mind. The following list of experiences will illustrate the types of psychological phenomena that might be classified under this heading: aesthetic appreciation (Platt, 1961); concept of self (Wylie, 1961); daydreams (Singer & Antrobus, 1963); night dreams (Dement & Wolpert, 1958); déja vu (Pickford, 1940); eidetic imagery (Haber & Haber, 1964); hallucinations (West, 1962); hypnagogic and hypnopompic imagery (McKellar & Simpson, 1954); the sudden awakening of reflective self-consciousness that has been called the 'I-am-me' experience (Spiegelberg, 1961); mnemonic systems and other structures (e.g. number forms), and strategies, (e.g. chess playing (Miller, Gallanter & Pribram, 1960)); naturally occurring hynotic-like experiences (As, 1962); nostalgia (McCann, 1943); organic sensations (Mason, 1959); peak experiences (Maslow, 1959); sensory imagery (Diehl & England, 1958); synaesthesia (Simpson & McKellar, 1955); reading fiction (Harding, 1962); religious experiences (James, 1902); revived emotions (Washburn, Deyo & Marks, 1924); and transcendent states induced through experimental meditation (Deikman, 1963).

These self-conscious states and processes of mind are not the same as such overt action patterns as aggressing, conforming, judging, helping, imitating or typewriting. Though many neo-behaviourists are concerned with mediational processes there has been relatively little concern with experiences as dependent variables to be described and accounted for in their own right.

It will be shown later in this paper that concern with the description of experiential states is not only of interest in its own right but may sometimes lead on to the development of important experiential constructs having the status of independent variables.

Subjective report

If experience constitutes a conceptually distinct class of psychological phenomena it will be necessary to reconsider the logical and empirical justification for the use of subjective report, which is its basic methodological procedure. Valuable discussions of this problem have been undertaken by Pratt (1939), Boring (1953), Bakan (1954), Zener & Gaffron (1962) and McKellar (1962) and these have been drawn upon in making the following comments.

In his book on the logic of modern psychology Pratt (1939) argued that 'introspection is merely another word for observation' (p. 56) and again that 'all phenomena are private to begin with, they become public only when their initial privacy is put into words. A rat in a maze can be as private as a pain in the stomach' (p. 99). At first sight this latter statement appears to be a paradox. Common sense leads us to make distinctions between private events and public events, or between subjective experiences and objective experiences. Some things, like 'rats in mazes' take place 'out there' beyond the surface of our bodies, other things like 'pains in the stomach' take place 'inside' our bodies and are only accessible to the inward eye of the body's owner.

This common-sense distinction, though convenient for some purposes, cannot be maintained as a logical distinction for use in a science of psychology. Whether I observe a rat in a maze or a pain in my stomach it is an observation that begins

with some sensory input and is mediated through me by the process of formulating a statement in words. In both instances 'I' have observed something and reported on it. At this stage it may or may not be accepted by others as an accurate observation. The usual requirements of reliability are check observations by the same person and/or by another independent observer. Such requirements can be satisfied in principle for subjective report as well as for objective report. The criterion of acceptability then, is that essentially the same report occurs under essentially the same conditions. This criterion does not require that the identical pain be observed by two observers, any more than it requires an identical rat in an identical maze to be observed by two observers. Observation of a rat in a Californian maze can be accepted as equivalent to observation of a similar rat in a similar West Australian maze providing that the total conditions are sufficiently similar to produce consistent reports.

From these total conditions the subset concerned with the specific conditions under which an observation is to be made are always of especial concern. The scientific psychologist wants to know exactly where, at what time, by what method and by whom the observation is reported. All attendant details of temperature, lighting, sounds, apparatus used, training of observers and so on are considered of fundamental importance to an assessment of scientific acceptability and it may be that the conditions under which an 'inside' observation can be observed and reported are more satisfactory than those that exist for a particular 'outside' observation. For example, the statement by a psychologist that after ingesting half a gramme of mescalin his head began to inflate like a football bladder and his legs to retract into his abdominal cavity may well be given more credence than the report of an amateur astronomer who says that he saw a flying saucer landing on the moon. The fact that one is the report of an inner experience and the other of an outer experience is not the criterion of its scientific acceptability. A scientific fact is a function of the reputability of the methods employed, including the reputability of the observer himself. Ultimately the reputability of the method is a function of the frequency with which it has been found to produce consistent results. When put in this form it becomes obvious that psychologists have been using subjective report ever since experimental psychology began.

Much of perception (Allport, 1955) and a good deal of the research by personality psychologists (Murray, 1958), social psychologists (MacLeod, 1951) and clinical psychologists (Rogers, 1958) involve a subject in the task of reporting on his subjective experience. Sometimes the subject is asked to report on experiences that occurred in the past, while at others, as in modern dream research, he is asked to report on an event that has only just occurred. Sometimes he is required to make a subjective report on some attribute of a simple experience and sometimes on a more complex one. Where the report is to be made on a simple experience from the recent past it is likely to be more reliable. In turn these more reliable reports enable the psychologist to set up and refine a more valid category system for the subject to use. One category system, of high generability, which has already demonstrated its validity is that of the semantic differential (Osgood, Suci & Tannenbaum, 1957; Osgood, 1962).

Experiential categories

Not only must the subject be helped to report accurately by greater clarity in defining the observational categories, but also the psychologist must develop reliable and valid dimensions or categories from the raw data with which he is provided. Two examples of this procedure will be considered.

A classical problem in psychology has been to provide an adequate definition of the stimulus. As Bakan (1956) observed 'It would be convenient to assume that the stimulus is whatever the experimenter defines it to be, but this assumption would oversimplify matters' (p. 369). One way of checking on assumptions of this kind is to employ specially designed post experimental questionnaires or interviews for use after every experiment which employs human subjects. The same recommendation has been made more recently and with much subtle elaboration by Orne (1962) in relation to the study of the demand characteristics of the situation in which the experimenter and his subjects meet.

One procedure by which relevant experiential variables might be created is outlined in Bakan's study in which 100 subjects were required to memorize a list of twenty nonsense syllables. These syllables were arranged on a blackboard in four columns of five and exposed for 5 min. Though this is hardly a typical procedure for the study of rote memory it had the function of high-lighting some unexpected but potentially significant independent variables that need to be controlled or systematically manipulated in any more sophisticated experiment.

A seventy-five item post-experimental questionnaire was answered immediately after the session. This was constructed on the basis of statements made by subjects in a pilot group on their methods and on their feelings and thoughts during the learning process, from a study of the literature on the topic of memorizing and finally from the experimenter's own analysis of what might be going on during the process of memorizing. Among the findings of this demonstration study were, that subjects who agreed with the following three statements recalled significantly more nonsense syllables than those who disagreed.

(1) 'I frequently would look away from the board or close my eyes and test myself by trying to see how many I knew'.

(2) 'I had the feeling that I was in a competitive situation and was trying to do better than the others.'

(3) 'I appreciate any break in the class routine.'

On the basis of this information it would be possible to develop a specific situational measure of task motivation and task rehearsal. Both these experientially derived variables could then be related to other personality and performance measures to determine whether more general dispositions were involved or whether they were important, but merely situation specific.

Bakan summarizes the main uses of these kinds of experiential data as '(*a*) to obtain a description of an experimental situation as experienced by the S, (*b*) to study the relationship between subjective and objective variables, (*c*) to discover significant variables, (*d*) to study relationships between subjective variables, and (*e*) to study the effect of experimental manipulations on subjective variables' (p. 378).

A second study serves to illustrate another way in which experiential data may be

used to form potentially significant psychological dimensions. In the field of imagery research, about which more will be said later, Gordon (1949, 1950) has developed a test for the measurement of imagery control. In its present form this test consists of a series of eleven situations that the subject must attempt to visualize. For example:

(1) Can you see a car standing in the road in front of a house?
(2) Can you see its colour?
(3) Can you see it in a different colour?

And so on to the last question which asks:

(11) Can you see the car all old and dismantled in a car cemetery?

In this test subjective reports are required on a set of preselected tasks of visual imaging. The job of the psychologist is to establish a way of categorizing or scoring these written answers so that the most discriminating kind of dependent or independent variable categories will result. Gordon (1950) found that by including only those persons in her well-controlled imagery category, who gave affirmative answers to each of the questions, that she obtained the greatest discrimination on a further task in which the rate of reversal on a Necker cube was measured. Subjects whose imagery was of the controlled type were found to have more control over the rate of reversal than subjects whose imagery was of the 'autonomous' or less well-controlled type.

In answering the eleven question on the Gordon Imagery Test the attempt is made to ensure that the same conditions hold as in any other test situation. First, the subject must understand the instructions and to this end it may be necessary to describe the nature of a visual image. If the subject is doubtful it may be a useful technique to ask him to fixate a black and white drawing for a brief period and then report what he sees when he closes his eyes or transfers his gaze to a sheet of off-white paper. This technique was used by Jaensch (1930) in his studies of eidetic imagery and might serve to acquaint a subject with one extreme kind of vivid visual imagery. Secondly, the subject must be honest in his endeavours to observe and report accurately. As with observers who are required to use other category systems one of the crucial determinants of reliability is the quantity and quality of training received. At present very little is known about the effects of training on the person who makes observations of his own subjective experiences.

To conclude this section on the problem of setting up adequate categories of dependent and independent variables from subjective reports two things should be noted. First, a great deal of important methodological information has accumulated in the area of content analysis which is relevant to the problem of coding and quantifying subjective reports (e.g., Allport, 1942; Berelson, 1952; McClelland, Atkinson, Clark & Lowell, 1953; Pool, 1959). Secondly, the procedure, though not the interpretation of its use, is in all essentials similar to the procedure used in constructing and validating a wide range of other psychological tests.

Theoretical implications

The question must now be raised as to whether the investigation of the variables of subjective experience has any general theoretical importance to psychology. There is a sense in which what one considers important is an individual matter of taste or of local and passing fashion. If we think of subjective experiences only as events to be

explained then it might still be argued that such problems should be left to those who are interested in them but can well be ignored by those who are not. If, on the other hand, it could be shown that there are any variables of subjective experience that affect a wide range of behaviour and experience then perhaps the theoretical importance of such independent variables would be more adequately established.

Already some grounds exist for believing that at least one variable of subjective experience will need to be incorporated within the framework of any systematic psychology. This experiential variable is imagery and some evidence is presented to show that it is of theoretical importance to the study of a wide range of psychological problems concerned with higher mental processes. In this regard the recent writings of Mowrer (1960) and the research interests of Piaget (cf. Gardner, 1962) are of great interest.

Imagery may be conceptualized as a class of cognitive abilities. As with all other abilities there is a wide range of individual differences in the capacity to form and control imagery in different sensory modes. Though there is some evidence of a general factor of imaging ability (Sutcliffe, 1962) there are three important sub-classifications that have been noted (e.g. Roe, 1951). People may be broadly categorized as visualizers, verbalizers or as mixed according to their preferred mode of thinking in different types of task. Perhaps the two major dimensions of imaging ability that are of importance are the vividness or clarity of the imagery and its controlability.

Among the major disputes relating to the concept of imagery in the past was its functional value or lack of it. Anti-imagists like Spearman (1930) maintained that imagery served no distinctively useful function, that a problem might be solved by employing imagery if one possessed it or equally well by utilizing inner speech or again without any conscious accompaniment at all, as in the imageless thought of the Würzburgers. Other psychologists like Davis (1932) sought to show that the possession of imagery was either essential or at least advantageous in the carrying through of particular cognitive and other tasks. He concluded an empirical investigation of his own with the pertinent comment that, 'If...the image experience...is shown to have behaviour consequences which are objectively demonstrable, then there is no longer any ground for dismissal of the image as epiphenomenal, unless the anti-imagist can demonstrate a correlation as close or closer, between the same objective behaviour differences and the mechanisms which he observes and which he claims are adequate for purposes of psychological description and explanation' (p. 657).

One of the tasks that is often thought to require vivid controlled visual imagery for its successful completion is the spatial manipulation of objects in imagination as when trying to imagine what a house will look like from drawings of its plan and elevation or what a room would look like when furniture and curtains are completely changed or rearranged.

In a study by Barratt (1953) twelve spatial tests were administered to a group of subjects and the resulting scores were intercorrelated and factor-analysed. Three principal factors were found: (I) Spatial manipulation, (II) reasoning, and (III) shape recognition. Subjective reports were obtained on the spontaneous imagery evoked during the performance of each test. Imagery, when it occurred, was found to be associated with the tests that had high loadings on factor I but to a lesser extent with

the tests that had high loading on factors II and III. 'The general conclusion, yet tentative, is that facility in imagery is important in spatial manipulation, is less important in shape recognition tasks and is unimportant in spatial analytic reasoning' (p. 162). It is interesting to note that in a study by Drewes (1958), in which an EEG measure of alpha blocking was used to differentiate visualizers from verbalizers, it was found that visualizers performed significantly better than verbalizers on the Guildford–Zimmerman Spatial Visualization test.

Studies of the type conducted by Barratt are of special significance because they help to make sense of earlier research which had produced conflicting results. Imagery appears to be of importance for high-level performance on one type of spatial task but not on all. Before the advent of factor analytic techniques a relatively unequivocal solution to this problem was not available.

One last illustration of the potential significance of imagery as a theoretical construct comes from the study of mental practice. Mental practice refers to the symbolic rehearsal of a physical activity in the absence of any gross muscular movements. When a golfer sits with eyes closed and in imagination goes through the motions of putting a golf ball he is engaged in mental practice. Since Sackett's (1934, 1935) early investigations of this phenomenon at least twenty-five further studies have been completed (Richardson, 1963). The trend in these studies provides support for the hypothesis that prior mental practice of a physical activity is associated with significant improvement in the actual performance of that activity.

A simple experiment by Vandell, Davis, & Clugston (1943) is an illustration, though in size of sample used and on other grounds it is not methodologically adequate. Three equated groups of Junior High School boys were used as subjects and the standard basket-ball free throw was the task. Group 1 who threw on the first and twentieth days with no practice of any kind in between showed 2 % improvement while group 2 who practised actual ball throwing on each of the eighteen days between first and twentieth days showed 41 % improvement. The members of group 3 were required to imagine themselves carrying out the throws on each of the eighteen intervening days and showed 43 % improvement.

Though much research needs to be carried out to determine the range of tasks, subjects and situational conditions which give rise to these improvements and the value of mixed patterns of mental and physical practice, enough evidence already exists to lend support to Perry's (1939) comment that, '...the implications for psychology and for education are far-reaching' (p. 5).

The possible importance of imagery as a factor in the efficient utilization of mental practice procedures has been suggested by Clark (1960) and others and has been subjected to empirical investigation in a pilot study by Start & Richardson (1964). These workers found that the greatest improvement in performance of a mentally practised gymnastic skill occurred in those subjects who possessed vivid controlled imagery, while the least improvement occurred in those who had vivid uncontrolled imagery. Those with weak controlled and weak uncontrolled imagery had mean scores in between. Though these results do not exclude the possibility of another covarying factor, such as capacity to concentrate, being responsible, there is at least a prima facie case for the role of imaging ability as a determinant of effective mental practice.

Conclusion

The purpose of this paper has been to explore some aspects of subjective experience as it exists in contemporary psychological research and theory. Attention has been drawn to the conceptual independence of experience and behaviour and to the logical and empirical justification for the use of subjective report. The procedure by which reliable and valid experiential categories might be constructed has been illustrated. The essential similarity of this method to that adopted in the construction of other inventories and tests is emphasized. Finally, the concept of imagery has been taken as an illustration of the possible theoretical significance of one class of experiential variables.

An earlier version of this paper was read to the West Australian group of the British Psychological Society (Australian Branch) at Perth, 1962.

References

Allport, F. H. (1955). *Theories of Perception and the Concept of Structure*. New York: Wiley.

Allport, G. W. (1942). The use of personal documents in psychological science. *Soc. Sci. Res. Coun. Bull.* no. **49**.

As, A. (1962). A factor analytic study of some subjective personal experiences and their bearing on theories of hypnosis. *Acta Psychol.* **20**, 196–209.

Bakan, D. (1954). A reconsideration of the problem of introspection. *Psychol. Bull.* **51**, 105–18.

Bakan, P. (1956). The collection and use of retrospective data. *J. Psychol.* **41**, 369–79.

Barratt, P. E. (1953). Imagery and thinking. *Aust. J. Psychol.***5**, 154–64.

Berelson, B. (1952). *Content Analysis in Communication Research*. New York: Free Press.

Boring, E. G. (1953). A history of introspection. *Psychol. Bull.* **50**, 169–89.

Clark, L. V. (1960). Effect of mental practice on the development of a certain motor skill. *Res. Quart.* **31**, 560–9.

Cronbach, J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psycho. Bull.* **52**, 281–302.

Davis, F. C. (1932). The functional significance of imagery differences. *J. exp. Psychol.* **15**, 630–61.

Deikman, A. J. (1963). Experimental meditation. *J. nerv. ment. Dis.* **136**, 329–43.

Dement, W. & Wolpert, E. (1958). The relation of eye movements, body motility and external stimuli to dream content. *J. exp. Psychol.* **55**, 543–53.

Diehl, C. F. & England, N. C. (1958). Mental imagery. *J. speech hear. Res.* **1**, 268–74.

Drewes, H. W. (1958). An experimental study of the relationship between electroencephalographic imagery variables and perceptual cognitive processes. Ph.D. thesis, Cornell University.

Gardner, R. W. (1962). Four discussions with Professor Piaget. *Bull. Menn. Clin.* **26**, 117–19.

Gordon, R. A. (1949). An investigation into some of the factors that favour the formation of stereotyped images. *Brit. J. Psychol.* **39**, 156–67.

Gordon, R. A. (1950). An experiment correlating the nature of imagery with performance on a test of reversal of perspective. *Brit. J. Psychol.* **41**, 63–7.

Haber, R. N. & Haber, R. B. (1964). Eidetic imagery. I: Frequency. *Percept. mot. Skills*, **19**, 131–8.

Harding, D. W. (1962). Psychological processes in the reading of fiction. *Brit. J. Aesthet.* **2**, 133–47.

Jaensch, E. R. (1930). *Eidetic Imagery*. London: Kegan Paul.

James, W. (1902). *The Varieties of Religious Experience*. New York: Longmans Green, 1929.

Koch, S. (1958). Epilogue In *Psychology: A Study of a science*, Vol. III. New York: McGraw-Hill.

McCann, W. H. (1943). Nostalgia: a descriptive and comparative study. *J. genet. Psychol.* **62**, 97–104.

McClelland, D. W., Atkinson, J. W., Clark, R. A. & Lowell, E. L. (1953). *The Achievement Motive*. New York: Appleton-Century-Crofts.

MacLeod, R. B. (1951). The place of phenomenological analysis in social psychological theory. In *Social Psychology at the Crossroads*, ed. J. J. Rohrer and M. Sherif. New York: Harper.

McKellar, P. & Simpson, L. (1954). Between wakefulness and sleep: hypnagogic imagery. *Brit. J. Psychol.* **45**, 266–76.

McKellar, P. (1962). The method of introspection. In *Theories of the Mind*, ed. J. M. Scher. New York: Free Press.

Maslow, A. H. (1959). Cognition of being in the peak experiences. *J. genet. Psychol.* **54**, 43–66.

Mason, R. E. (1959). Three studies relating internal sensory experiences and feeling reactions. *J. gen. Psychol.* **60**, 211–28.

Miller, G. A., Gallanter, E. & Pribram, K. H. (1960). *Plans and the Structure of Behaviour*. New York: Holt.

Mowrer, O. H. (1960). *Learning Theory and the Symbolic Processes*. New York: Wiley.

Murray, H. A. (1958). Preparations for the scaffold of a comprehensive system. In *Psychology: A Study of a Science*, vol. III, ed. S. Koch. New York: McGraw-Hill.

Orne, M. T. (1962). On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *Amer. Psychologist*, **17**, 776–83.

Osgood, C. E., Suci, J. & Tannenbaum, P. H. (1957). *The Measurement of Meaning*. Urbana: University Illinois.

Osgood, C. E. (1962). Studies on the generality of affective meaning systems. *Amer. Psychologist*, **17**, 10–28.

Perry, H. M. (1939). The relative efficiency of actual and imaginary practice in five selected tasks. *Arch. Psychol.* **34**, 5–75.

Pickford, R. W. (1940). Three related experiences of déjà vu. *J. Person.* **9**, 152–9.

Platt, J. R. (1961). Beauty: pattern and change. In *Functions of Varied Experience*, ed. D. W. Fiske and S. R. Maddi, Homewood: Dorsey.

Pool, Ithiel de Sola (ed.) (1959). *Trends in Content Analysis*. Urbana: University Illinois.

Pratt, C. C. (1939). *The Logic of Modern Psychology*. New York: MacMillan.

Richardson, A. (1963). Mental practice: a review and discussion. *Bull. Brit. psychol. Soc.* **16**, no. 51, 28a–29a.

Roe, A. (1951). A study of imagery in research scientists. *J. Personal.* **19**, 459–76.

Rogers, C. (1958). A theory of therapy, personality and interpersonal relationships, as developed in the client centred framework. In *Psychology: a Study of a Science*, vol. III, ed. S. Koch. New York: McGraw-Hill.

Sacket, R. S. (1934). The influence of symbolic rehearsal upon the retention of a maze habit. *J. gen.* **10**, 376–95.

Sackett, R. S. (1935). The relationship between amount of symbolic rehearsal and retention of a maze habit. *J. gen. Psychol.* **13**, 113–28.

Simpson, L. & McKellar, P. (1955). Types of synaesthesia. *J. ment. Sci.* **100**, 141–7.

Singer, J. L. & Antrobus, J. S. (1963). A factor analytic study of day dreaming and conceptually related cognitive and personality variables. *Percept. mot. Skills*, **17**, 187–209.

Spearman, C. E. (1930). *Creative Mind*. London: Nisbet.

Spiegelberg, H. (1961). On the 'I-am-me' experience in childhood and adolescence. *Psychologia*, **4**, 135–46.

Start, K. B. & Richardson, A. (1965). Imagery and mental practice. *Brit. J. educ. Psychol.* **34**, 280–4.

Sutcliffe, J. P. (1962). The relation of imagery and fantasy to hypnosis. Progress report on N.I.M.H., Project M-3950, University Sydney, Australia, 1962.

Vandell, R. A., Davis, R. A. & Clugston, H. (1943). The function of mental practice in the acquisition of motor skills. *J. gen. Psychol.* 1943, **29**, 243–50.

Washburn, M. F., Deyo, D. & Marks, D. (1924). A further study of revived emotions. *Amer. J. Psychol.* **25**, 113–20.

West, L. J. (ed.) (1962). *Hallucinations*. New York: Grune and Stratton.

Wylie, R. C. (1961). *The Self Concept*. Lincoln: University Nebraska Press.

Zener, K. & Gaffron, M. (1962). Perceptual experience. In *Psychology: A Study of a Science*, vol. iv, ed. S. Koch. New York: McGraw-Hill.

(*Manuscript received* 15 *June* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 233–242
Printed in Great Britain

PROBABILITY LEARNING IN STEP-INPUT TRACKING*

By C. B. GIBBS

Defence Research Medical Laboratories, Toronto, Canada

Six men and six women tracked stimuli that demanded responses of unequal probability. The control-display relation was directionally incompatible. Half the subjects used their non-preferred hand. Many large directional errors occurred in early practice and these were amended after a mean delay of 0·24 sec. With continued practice, small errors persisted mainly in responses of low probability but the mean amendment time fell to 0·11 sec. These errors provided new, highly sensitive measures that revealed differences in performance associated with sex, hand preference and probability ($P < 0{\cdot}01$).

The results are compatible with hypotheses that the speed, direction and extent of movement are determined by negative proprioceptive feedback and integral-error control (Gibbs, 1954).

There are lawful relations between the direction of movement and the discharge of specific groups of proprioceptors (Mountcastle, Poggio & Werner, 1963). The speed of movement is closely related to the rate of change of frequency in the discharge of primary nerve endings (Matthews, 1933). The speed and direction of movement could therefore be controlled directly by proprioceptive feedback, once the relevant relations were learned. The extent of movement could be determined by integrating the speed signals from proprioceptors over time, to provide integral-error control (Gibbs, 1954). It is hypothesized that new movements are first controlled by exteroceptors, but that detailed duties of monitoring are delegated to proprioceptors for the longest possible period. The degree of dependence on vision depends on the probability and predictability of the outcome of specific responses. Delegation releases exteroceptors and the limited span of attention from detailed duties of monitoring.

A hypothetical function of proprioception is to provide negative feedback which ensures that a movement initiated by an error-stimulus reduces that error. Negative feedback depends on a definite directional relation between input data, and the feedback data which define the output. A reversal of the normal learned relation between visual and proprioceptive data should therefore lead to positive feedback by which response to an error-stimulus increases the error.

Two distinct types of learning underlie the effective control of movement. The learning of lawful relations between movements and sensory feedback (output-feedback relations) is to be distinguished from the learning of serial, S–R probability relations; they are exemplified by the tracking task that is described. A major purpose of the experiment is to demonstrate the existence, and some major effects of both types of learning.

The above hypotheses lead to unequivocal predictions and to consistent explanations of known phenomena. In a task of step-input tracking the direction of successive responses is not always equally probable; such tasks are convenient for the study of serial probability relations. It is easy to provide an incompatible directional relation between control and display, i.e. a control/display relation where the control

* Defence Research Medical Laboratories Research Paper no. 541, PCC No. D 77–94–20–23.

joystick and the display cursor move in opposite directions. Such a relation is well adapted for the demonstration of positive feedback in early practice, which results in large directional errors and considerable delays before errors are amended. By hypothesis, the unfamiliar but constant control/display relation can be learned and negative feedback again develops with a consequent reduction of errors in responses of both high and low probability.

The delegation of monitoring from vision to proprioception, or from conscious to automatic level, reduces the time taken to amend directional errors. The approximate reduction should be from 0·25 to 0·10 sec; the former time represents visual reaction time, and the latter corresponds to known latency in the motor-proprioceptor circle of nerves (cf. Ruch, 1951). Amendment times of 0·10 sec would contrast with findings of psychological refractoriness, e.g. that in responses to two closely spaced stimuli, the reaction time to the second is longer than that to the first (e.g. Vince, 1948; Craik, 1948; Hick, 1948; Welford, 1952, 1959; Davis, 1956, 1957, 1959; Adams, 1961, 1962). Other predictions from the hypotheses are that the number of directional errors and the duration of visual reaction times depend on the relative probabilities of responses. The data reported here are compatible with hypotheses.

Method

Six men and six women of ages ranging from 23 to 54 years, used a joystick to position a cursor on an oscilloscope and track a target spot of light. Movement was limited to the horizontal plane. Three subjects of each sex used their preferred hand; the others used their non-preferred hand. In the main study the control/display relation was incompatible, but a control group of six subjects used a compatible relation in an otherwise identical procedure. Subjects were instructed to respond as rapidly and accurately as possible.

The target appeared for 1 sec in any of five different positions, 1·125 in. apart, disappeared for 1 sec and reappeared in a new position. Each position and each possible pair of successive positions was used equally frequently in one complete 'run' of 100 steps. There was a three to one probability that a target at position 2 would move to the right rather than outward to position 1; a target at position 4 was more likely to move left than right. Steps are termed 'probable' when the movement actually demanded conformed to the higher probability; steps outward from 2 or 4 are termed 'improbable'. Responses beginning at position 3 are called 'equi-probable' because the two possible directions of movement were equally probable. Movements from positions 1 or 5 were 'unequivocal' with respect to direction.

Records of target and joystick movements were analysed for errors in the initial direction of movement, for response latencies and for amendment times. Response latency is the interval between the onset of a stimulus and the beginning of a response; the definition applies also to reaction time but a different term is needed for the delays which arise in tracking tasks. Amendment time is the interval between the beginning of an incorrect response and the commencement of an amended movement.

Results

The control group using a compatible control/display relation made 8 errors in 600 responses as compared with the 272 errors made in the 1200 responses of the experimental group. There was a small but significant ($P < 0{\cdot}01$) difference of 0·04 sec in the mean response latencies of all types of responses, which also favoured the compatible control/display relation. Table 1 shows the proportion of experimental group errors in responses of different probability at the three stages of practice indicated in column 1. The first entry in column 2 under the heading E/R shows that between steps 1 and 33 the group made 37 errors (E) in a total of 48 improbable responses (R).

The adjacent entry expresses the fraction as a percentage (77%). Other entries in the table show similar relations for other responses. The main findings are that the great majority of errors occurred on the relatively infrequent improbable and equiprobable responses. There was a monotonic decrease in all error percentages as practice continued, except for an increase on equiprobable steps in the last stage of practice. All twelve subjects made errors on step 1 which demanded an equiprobable response, but following step 33 a maximum of six subjects made errors on these steps. Nine subjects made errors on the first and on the last appearance of improbable steps (step 16 and step 89) percentage errors remained high (60%) in the last stage of practice. In contrast, errors virtually disappeared in unequivocal responses.

Table 1. *Total number of errors (E) in responses (R) of different probability at different stages of practice (steps* 1–33, 34–66 *and* 67–100)

		Improbable	Equiprobable	Probable	Unequivocal
	(1) Steps	(2) E/R %	(3) E/R %	(4) E/R %	(5) E/R %
	1–33	37/48 77	44/72 61	38/120 32	18/156 12
	34–66	25/36 69	27/96 28	13/120 11	4/144 3
	67–100	22/36 60	29/72 40	12/120 10	1/180 0·6
Totals	1–100	84/120 70	100/240 42	63/360 18	23/480 5

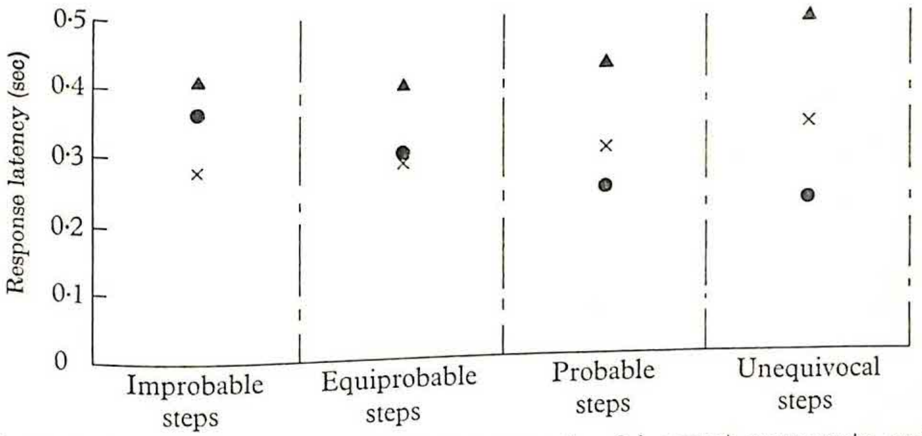


Fig. 1. Relation between response latencies for correct and incorrect movements, and corrective response latencies, on steps of different probability ($n = 12$). ●, Mean response latency for correct responses. ×, Mean response latency for incorrect responses ▲, corrective response latency (response + amendment).

Response latencies, like errors, were markedly reduced in early practice. The mean response latencies shown in Fig. 1 relate to steps 15 to 100; the stage where response latencies became reasonably consistent. Fig. 1 shows the striking and significant ($P < 0{\cdot}01$) monotonic relation between serial probability and the mean response latencies of correct responses; a relation similar to that noted by Hyman (1953). The figure shows that on improbable steps the response latency for a correct movement was nearly 0·10 sec longer than the response latency for an error. This is the expected relation between accuracy and speed. It was very surprising that the response latencies of correct and incorrect responses were virtually equal on equiprobable steps

and that the latter response latency actually exceeded the former on probable and unequivocal steps, i.e. the larger latencies produced more errors.

The correlation between the response latencies and the errors of subjects was not significant, i.e. the subjects who responded most rapidly did not make most errors. Three male subjects had shorter response latencies and fewer errors than the group mean, and provided the three best combined scores. One male subject made the poorest combined score. The range of individual errors was 10–37.

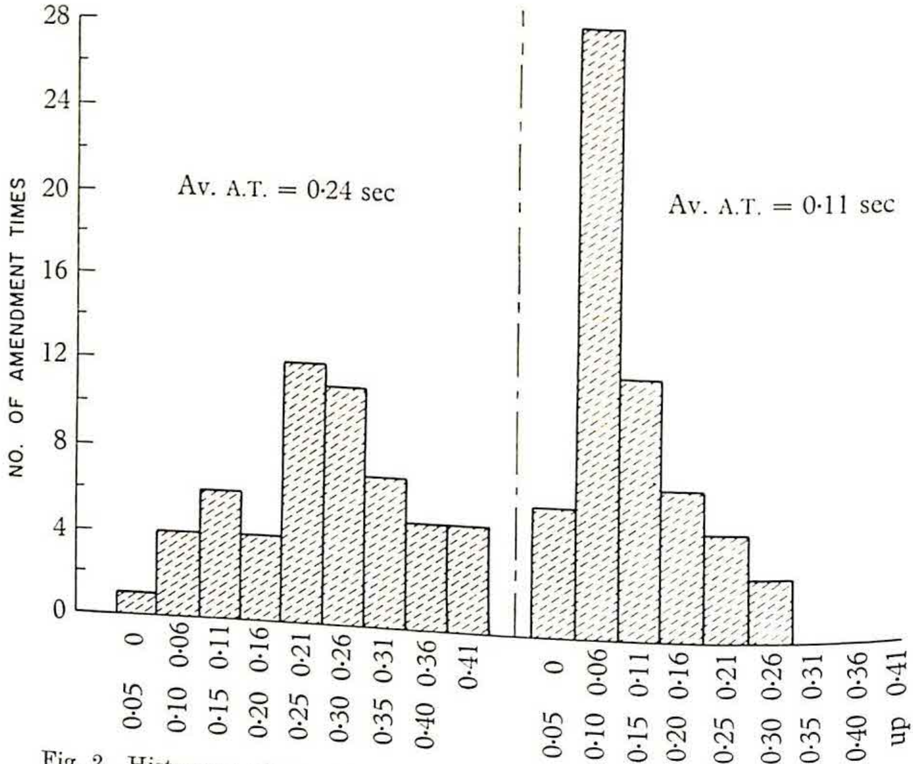


Fig. 2. Histogram of amendment times in steps 1–10, and steps 71–100.

A major prediction from hypotheses was that amendment times would be reduced from about 0·25 to 0·10 sec. Fig. 2 shows the actual distribution of amendment times in early and late practice (steps 1–10 and steps 71–100). There is a striking and significant change in mean amendment time from 0·24 to 0·11 sec, as predicted.

The top half of Fig. 3 shows the responses of a typical female subject on steps 15–20. The square form steps represent the various positions of the target from 1 to 5, denoted respectively by the squares of minimum and maximum height. The thick base lines between steps of different heights represent the 1 sec periods of target absence. The continuous line above the steps is the subject's tracking response. A correct response was always in a direction that matched the changed height of a new step. The first step demanded an improbable response from position two to one; the tracking line should have moved downward in the figure. In fact, the subject made a small anticipatory movement in the probable and incorrect direction before light one appeared. Following its appearance, the subject made a large, all-or-none response of approximately correct extent in the wrong direction. Anticipatory movements were common in the subjects who made most errors. The response from

position 2 to 4 was initially in the correct direction but was reversed, approximately 0·10 sec after movement began. The reversal caused a large error in direction, which was amended after a delay that was compatible with visual reaction time. Responses that were initially correct, but rapidly reversed, accounted for 12 % of all the errors made in steps 1 to 33. Errors of less than 0·125 in. at the display would not be detected in analysis because of size reduction in recording. The lower half of Fig. 3 shows responses at a later stage in practice (steps 51–56). Directional errors occurred in making a probable response from position 4 to 2 and an improbable response from position 2 to 1. In general, the size of errors was much reduced by practice.

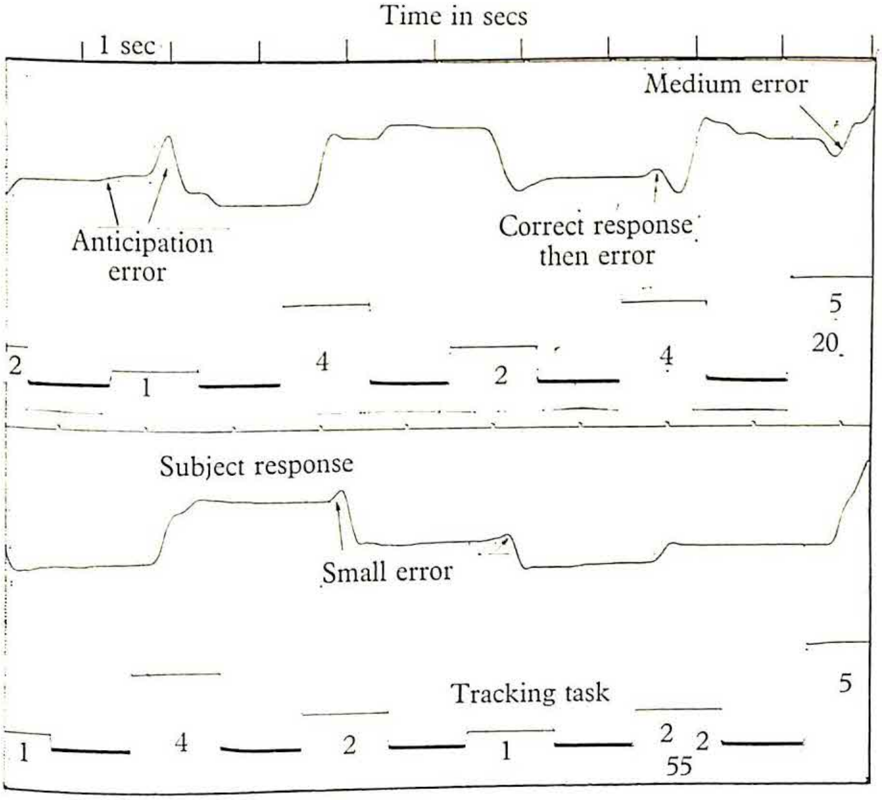


Fig. 3. Response of typical woman subject on steps 15–20 (upper) and steps 51–56) (lower).

In the 1200 responses made by the experimental group, 272 were initially in the wrong direction. The women made 168 and the men made 104 errors, a significant difference ($P < 0{\cdot}01$). Five women and one man produced the six highest error scores. The errors made by both sexes were distributed in a similar manner between responses, of different probabilities, indicating similar estimates of response probability. Without regard to sex, 158 errors were made by the six subjects who used their non-preferred hand and 114 errors were made by the other six subjects who used their preferred hand. This difference also was significant ($P < 0{\cdot}01$). As noted above, the control group using a compatible control/display relation made only eight errors in 600 responses.

Discussion

The small, rapidly amended errors observed in these tests have not been reported previously. Errors and response latencies provided highly sensitive measures that emphasized small but significant differences in performance due to hand, probability and sex.

It is not yet possible to explain why the majority of women subjects made more errors than most of the men. There was no reliable sex difference in probability matching; contrary to some established prejudices, the difference in errors did not stem from a feminine tendency to repose undue confidence in highly improbable events, e.g. that a target starting from the edge of the oscilloscope would move off the screen entirely.

Errors increased sharply when subjects were placed under minor stress, either by responding to stimuli of low probability or by using their non-preferred hand. These results indicate that these sensitive measures of performance may be particularly useful in studying stress. The method has now been applied successfully to a study of effects of alcohol and a report is in preparation.

The problem of control

Lawful relations must exist in any system to permit effective control of the speed, extent and direction of movement. In organisms, practice can improve performance only to the point where these relations are learned. There is a definite relation between the direction of movement and the discharge of specific groups of proprioceptors (Mountcastle *et al.* 1963). High correlations exist between the rate of movement and the rate of change of frequency of the kinaesthetic discharge (Matthews, 1933). Proprioception could therefore monitor direction and speed directly; speed could be integrated over time in order to determine the extent of movement by integral error control.

Correlations also exist between movement and visual feedback but the typical errors of early practice show that vision does not continuously monitor rapid movements. Many movements were almost correct in extent but ran in the wrong direction for about 0·25 sec. During that period visual feedback was not effective in amending error and it cannot therefore be credited with the continuous monitoring of direction, speed or extent. Many early responses were initially correct but were reversed after approximately 0·10 sec to produce a large directional error. Had visual feedback been effective in that time, the correct initial adjustment would have continued.

Data reported by Helson & Steger (1962) suggest that visual feedback can produce a motor effect in 0·1 sec. In the transition from visual to proprioceptive monitoring, some rapid amendments were possibly based on vision, but the purpose of delegation is to relieve visual attention of uneconomic duties of detailed monitoring.

Probability, latency and error

The delegation of monitoring from vision to proprioception, effects a change in the time for amending errors, from about 0·25 to 0·10 sec. The former period represents visual reaction time, the latter is the known latency in the motor-proprioceptor circle

of nerves (Ruch, 1951) and the minimum time to amend responses that are not monitored by vision, e.g. hand-tapping (Dresslar, 1892).

In early practice, the mean time to amend errors was 0·24 sec (Fig. 2). Subjects expected that a rapid, primary response could run for that period, under proprioceptive control, without developing serious error requiring a visual check. The expectation was invalidated by the reversal of normal directional relations, which permit control by negative, proprioceptive feedback. Subjects were presented with a choice between speed and accuracy until the new relation was learned. They could regress to slow, visual monitoring which would minimize errors, but increase the time needed to acquire the target, which appeared only briefly. Alternatively, they could respond rapidly using proprioceptive feedback, but the reversed relation would then produce frequent directional errors.

The mean amendment time was 0·11 sec in late practice, indicating that subjects were probably using proprioceptive monitoring, despite a high percentage of errors on the less probable steps. The subjects were not aware of these rapid amendments. Delegation therefore largely relieved visual attention of the duties of detailed monitoring, although the incompatible relation was not fully learned. The overall error percentage in the last stage of practice was about 16 %, and only 1·3 % of errors were made by the control group using a compatible relation.

A clear distinction exists between the lawful relations of movement and feedback, and the serial probability relations between input and output, that are exemplified by the different types of tracking steps. The latter permit tests of the hypothesis that the degree of visual attention depends on the input-output probability relation. Specifically, highly probable responses are initiated with a minimum of visual attention, which is indicated by short, visual response latency. Fig. 1 indicates that the latency is related directly to response probability. Table 1 shows that the largest proportion of errors occurred on the less probable steps, so that a further disproportionate increase of visual attention would be needed on these steps, to minimize the errors.

The stimuli appeared at equal time intervals but there was always uncertainty of the extent of impending responses. The latency of a correct response was 0·23 sec on unequivocal steps, where the demand for direction was certain, but each of four different extents was equally probable. In steps starting from positions 2 or 4, the increase of directional uncertainty raised latency to 0·27 sec for a correct response in the probable direction, but a similar mean latency of 0·28 sec produced errors on directionally improbable steps. It was necessary to increase latency to 0·37 sec to obtain correct responses on the latter steps. There were three times as many probable as improbable steps. Subjects therefore had a choice between short latency, producing about 25 % of errors, and a long latency on both probable and improbable steps, which would minimize errors. In late practice, errors could be amended in 0·10 sec and this was near the difference between the short and long latency. Time loss would therefore be negligible even when errors were made, and time was important owing to brief target presentation. A strategy that minimized the time to acquire the target, rather than error, would produce 25 % of errors on steps starting from positions 2 or 4. In late practice, the average group errors were 22 %.

It would be expected that increased care in identifying the direction demanded

would increase response latency and diminish error. This relation obtained on improbable steps, but on equiprobable steps the latencies of correct and incorrect responses were virtually equal at 0·30 and 0·29 sec respectively. On these steps, 50 % of responses would be correct by chance, even with the 'short' latency of about 0·30 sec, and there would be little time loss in amending errors, as compared with using the 'long' latency of 0·37 sec for all responses. Again, the group minimized target acquisition time rather than error, and accepted 40 % of error in late practice.

Fig. 1 shows that on probable and unequivocal steps, errors were actually associated with the longer latencies, i.e. the opposite of the expected relation. However, most of the relevant, plotted data were obtained in early practice, when error and long latency were associated on all types of responses. After step 67, there were only 13 errors on probable and unequivocal steps; these were probably due to temporary regressions in learning the incompatible, control/display relation. They are not significant exceptions to the normal, inverse relation between accuracy and speed.

The development of a short amendment time of 0·1 sec was an alternative to accepting long response latency on all steps, and it produced an effective compromise between the opposed requirements for accuracy and speed. The relations that developed, between response probability, latency and error reflect surprisingly high predictive capacity. Subjects developed, in 1 min practice, a highly adaptive strategy based upon relations that are extremely complex and obscure at the conscious, intellectual level. The group as a whole, and the women in particular, adopted the strategy that minimized target acquisition time, rather than error. The verbal reports of subjects revealed no awareness of rapid amendments; control and computations were effected at subconscious level.

Fig. 1 illustrates the significant finding that uncertainty of direction had a greater effect on the latency of correct responses, than doubt of extent. For example, four different extents could be demanded on a directionally unequivocal step, but only one extent was involved in a directionally improbable step. 'Correct' response latencies were 0·37 and 0·23 sec respectively.

Psychological refractoriness

Data on the psychological refractory period are contained in the studies cited previously. In responses to two closely spaced, unpredicted stimuli, the latency of the second response is usually longer than that of the first. A decision to respond cannot be revoked immediately by a prompt indication of error. The amendment times of the present study were, in effect, the response latencies to the second of two closely spaced stimuli. In early practice, directional errors persisted for about 0·25 sec and demonstrated the typical 'all-or-none' refractory effect, but the amendment time was reduced to 0·10 sec in late practice.

By previous hypothesis, the reduction of delay is based on learned, serial probability relations which were not apparent initially. The subjects could learn the relations between step probability, response latency and errors; for example, that a latency of 0·28 sec, on steps starting from positions 2 or 4, would produce an error percentage of 25 %. To this extent, errors became predictable. Once a response was initiated, the direction of any error was determined, and time of detection would be

set by a relatively small and uniform perceptual delay. There was no uncertainty of appropriate response when the error signal occurred; hence, visual attention and central decision were not involved in amendments.

The data show that one type of refractoriness can be reduced by practice. Learning permits prediction of the results of responses, and the delegation of monitoring from visual attention to proprioceptive mechanisms which function at subconscious level.

The alternative view

Whitteridge (1960) provides a concise, impartial account of the historic controversy between protagonists of the inflow and outflow theories of movement control. In the former, precise, directed movements cannot be made without either exteroceptive or proprioceptive information (feedback), on the states of the controlled member. In the outflow view, proprioception has no important, central functions in control. Practice establishes learned patterns of motor impulses in the brain, and each is appropriate to nullify an error of specific extent. The incompatible relation could be learned, by the incremental, automatic reinforcement of 'correct' motor volleys, and the progressive inhibition of 'incorrect' patterns. At some stage of practice, conflicting volleys of motor impulses could be despatched over the direct and the indirect motor pathways to muscle (Eldred, Granit & Merton, 1953). Known difference in conduction time could produce rapid reversal of movements, and so account for the 0·1 sec interval between successive movements, without invoking notions of feedback.

The proposed, differential reinforcement necessitates feedback, in order to discriminate between correct and incorrect responses, but the converse does not hold. The essence of inflow theory is that patterns of proprioceptive sensations, rather than motor impulses, form the content of learning. Notions of motor reinforcement are, at best, redundant, and have nothing to commend them except respectable antiquity. The hypothesis of automatic reinforcement implies a 'stamping in' process to ensure that the appearance of a familiar stimulus triggers its associated response. No provision is made for adaptive variability, demonstrated by subjects' compromise between the opposed needs to acquire the target quickly and to minimize error. In the outflow view, the appearance of either an improbable step from position 4 to 5, or an unequivocal step from 1 to 2 would trigger the same pattern of motor impulses. Hence, errors and response latency would be equal on steps of similar extent, in the same direction, irrespective of their probability. The data flatly contradict the predictions based on outflow theory. The subjects were undoubtedly learning the new directional relation between output and feedback, and the input-output, serial probability relations that were present in the task.

Acknowledgement is due to Mr O. Logan who provided valuable help in conducting the tests and in analysing data.

References

Adams, J. A. (1961). Human tracking behaviour. *Psychol. Bull.* **58**, 55–9.

Adams, J. A. (1962). Test of the hypothesis of psychological refractory period. *J. exp. Psychol.* **64**, no. 3, 280–7.

Craik, K. J. W. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.* **38**, 142–8.

DAVIS, R. (1956). The limits of the 'psychological refractory period'. *Quart. J. exp. Psychol.* **8**, 24–38.

DAVIS, R. (1957). The human operator as a single channel information system. *Quart. J. exp. Psychol.* **9**, 119–29.

DAVIS, R. (1959). The role of 'attention' in the psychological refractory period. *Quart. J. exp. Psychol.* **11**, 211–20.

DRESSLAR, F. B. (1892). Some influences which affect the rapidity of voluntary movements. *Amer. J. Psychol.* **4**, 514–27.

ELDRED, E., GRANIT, R. & MERTON, P. A. (1953). Supraspinal control of the muscle spindle and its significance. *J. Physiol.* **122**, 498–523.

GIBBS, C. B. (1954). The continuous regulation of skilled response by kinaesthetic feedback. *Brit. J. Psychol.* **45**, 24–39.

HELSON, H. & STEGER, J. A. (1962). On the inhibitory effects of a second stimulus following the primary stimulus to react. *J. exp. Psychol.* **64**, no. 4, 201–5.

HICK, W. E. (1948). The discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.* **1**, 36–44.

HYMAN, R. Stimulus information as a determinant of reaction time. *J. exp. Psychol.* **45**, 188–96.

MATTHEWS, B. H. C. (1933). Nerve endings in mammalian muscle. *J. Physiol.* **78**, no. 1, 1–53.

MOUNTCASTLE, V. B., POGGIO, G. F. & WERNER, G. (1963). The relation of thalamic cell response to peripheral stimuli varied over an intensive continuum. *J. Neurophysiol.* **26**, 807–34.

RUCH, T. C. (1951). Motor Systems. In *Handbook of Experimental Psychology*, (ed. Stevens, S. S.) New York: Wiley.

VINCE, M. A. (1948). The intermittancy of control movements and the psychological refractory period. *Brit. J. Psychol.* **38**, 149–57.

WELFORD, A. T. (1952). The psychological refractory period, and the timing of high-speed performance: a review and a theory. *Brit. J. Psychol.* **43**, 2–19.

WELFORD, A. T. (1959). Evidence of a single channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.* **11**, 193–210.

WHITTERIDGE, D. (1960). Central control of eye-movements. Chapter 42 in *Handbook of Physiology*, Vol. II: *Neurophysiology* (ed. Magoun, H. W.). Washington: American Physiological Society.

(*Manuscript received* 31 *July* 1964)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 243–254
Printed in Great Britain

ANALYSIS OF VISUAL AND PROPRIOCEPTIVE COMPONENTS OF MOTOR SKILL BY MEANS OF A DRUG

BY D. LEGGE
University College London

Skilled manual responses depend upon information about the position of the hand which is to be moved. In order to throw light on the way in which CNS depressant drugs impair skill, an experiment was performed to study the effect of nitrous oxide on the perception of hand position by vision and by proprioception. The results show that these two modalities were less efficient in combination than was either separately. The drug increased the variability of performance irrespective of the perceptual conditions. The drug also produced systematic changes in constant error, by its action either on vision or on proprioception. Drug-induced *increases* in the size of handwriting may be explained as changes which compensate for the effects of the drug on perception.

Drugs which have predominantly depressant effects on the central nervous system usually impair the performance of complex perceptual-motor skills. This effect has been experimentally demonstrated with several depressant drugs including alcohol (Jellinek & McFarland, 1940; Drew, Colquhoun & Long, 1959; Carpenter, 1962), barbiturates (Payne & Moore, 1955; Kornetsky, Humphries & Evarts, 1957) and nitrous oxide (Steinberg, 1954; Legge, Steinberg & Summerfield, 1964). The complex skills studied in these experiments were of various kinds including a simulated driving task, a multi-dimensional compensatory tracking task and handwriting. There has been relatively little systematic analysis, however, of components of these skills which may be particularly sensitive to the effects of depressant drugs.

The execution of complex skills must depend upon the ability to position an effector, like the hand, accurately in space. Although complex skills involve considerably more than this simple positioning ability (Bartlett, 1948; Poulton, 1957) their execution would be impossible without it. By determining the effects of depressant drugs on a simple positioning response it is possible to begin to analyse the effects of depressant drugs on more complex skills, and the skills themselves.

The accuracy of rapid hand positioning depends upon two major factors: first, the accuracy with which the position of the target and the initial position of the hand are localized and, secondly, the adequacy of the *implicit* response which must be prepared in order to move the hand. It is important to differentiate between these two factors. A drug which decreases the accuracy of hand positioning responses may do so either by altering the perception of spatial location or by acting on the processes underlying the organization and production of a response.

In some skills positional information about the target and about the hand may be mediated by the same modality, usually vision. However, one of the characteristics of increased skill at a task is the transfer of some of the monitoring functions of vision to proprioception (Fitts, 1951; Fleishman & Rich, 1963). In fact some skills, including many ball games, cannot be performed if vision is used to locate the effectors because vision is completely occupied with the target. The importance of proprioception is illustrated by the plight of patients with *tabes dorsalis*, which results from

destruction of the dorsal roots of the spinal cord that convey proprioceptive information. A tabetic patient has difficulty in walking unless allowed to look at his feet, and if the upper spinal roots are involved his hand movements are much more clumsy than usual (Gibbs, 1954). It seems likely that in the normal execution of a skilled hand positioning response, proprioception provides information about the initial position of the hand and that vision provides information about the location of the target. The organization of an adequate positioning response must then depend upon a comparison of these two sets of information which are mediated by different modalities.

The experiment reported here was carried out to investigate the selective effects of several doses of a central nervous system depressant drug, the anaesthetic gas, nitrous oxide, on the perception of hand position and on the way in which visual and proprioceptive information may be integrated.

The aim was to investigate the effects of small doses of nitrous oxide, on the precision and accuracy with which the position of the hand is perceived, and in particular, to determine any differential effects of nitrous oxide on proprioception and vision. The accuracy with which subjects could align a pointer with a stationary target line was assessed under different perceptual conditions and under different doses of nitrous oxide. The effects of the same doses of nitrous oxide were also observed on handwriting which is a complex over-learned perceptual-motor skill.

METHOD

In the alignment task subjects had to move the point of a vertical stylus, controlled by the right hand, to a position in line with a horizontal target line slightly to the left of and perpendicular to the forward line of sight. There was a gap between the target line and the stylus so that subjects had to position the stylus on an imaginary extension of the target line rather than upon the target line itself.

The target line could be perceived either by binocular vision or by proprioception. The stylus controlled by the subject could also be perceived either visually or proprioceptively. Therefore there were four different combinations of modalities and each combination represented a different perceptual condition. The four perceptual conditions were as follows: both target and stylus perceived visually (VV), target perceived visually and stylus proprioceptively (VP), target perceived proprioceptively and stylus visually (PV), and both target and stylus perceived proprioceptively (PP).

Three doses of nitrous oxide were used: 10, 20 and 30% nitrous oxide in oxygen, with air as control. Nitrous oxide was chosen for several reasons. It is a gas which depresses activity in the central nervous system. It is simple to administer, its effects occur relatively quickly, its concentration may be easily controlled and subjects quickly recover after the experiment. The largest dose of nitrous oxide given in this experiment was very much smaller than the doses normally required to induce anaesthesia (Goodman & Gilman, 1955) but was probably large enough to produce some analgesia (Henrie, Duncan & Home, 1960).

Combination of the four perceptual conditions with the four doses of nitrous oxide makes a total of sixteen experimental treatments. In order to reduce the effect of individual differences in skill and in response to the drug, each subject performed the alignment task under each of the experimental treatments. A counter-balanced latin square design was used which had the special property that each treatment preceded each other treatment (Williams, 1949). In this way both order effects and the residual effects of a preceding treatment were balanced. Each replication of the design required four subjects. Two replications were made.

Subjects

The eight subjects who took part in the experiment were right-handed male student volunteers. Before taking part subjects were screened for histories of physical or mental abnormality. Each subject spent about four hours in the experimental situation and was paid for his services.

Apparatus

The 6 in. target line was set in the horizontal plane 20 in. below eye level. The 'line' was a sliver of chemically blackened brass set in a sheet of sand-blasted white opal Perspex. The distance of the target from the subject varied from trial to trial about an average of 12 in. The stylus was a vertical steel rod of similar dimensions and shape as a pencil. The rod was suspended rigidly from a gantry which restricted its movement to a line parallel to the subject's forward line of sight and at right angles to the target line. The distance between the target and the line of movement of the stylus was $2\frac{1}{2}$ in. The whole of the working area was evenly illuminated. When the subject was required to perceive the location of the target proprioceptively the line was replaced by a metal strip 6 in. long and $\frac{1}{8}$ in. high. When the stylus was perceived visually the subject controlled its movement indirectly through a displaced lever so that he did not also receive proprioceptive information about its location. Restriction of vision was achieved by inserting screens. The position of the stylus was displayed by projecting a magnified image of a rule fixed to the stylus gantry, on to a screen outside the subject's field of vision. By incorporating a vernier scale the position of the stylus could be measured to the nearest 0·1 mm.

Nitrous oxide/oxygen mixtures and air were administered to subjects from a modified anaesthetic trolley through a face mask. All gases were scented with lavender oil. The technique has been described in detail by Steinberg (1954).

Procedure

Each subject attended at the same time on each of four consecutive days. On each day the subject carried out the alignment task under each of the four perceptual conditions, while being given one concentration of nitrous oxide throughout the session. Five minutes were allowed for equilibration (cf. Kety & Schmidt, 1948). During this period subjects copied in pencil digits taken from a table of random numbers. The digits were typewritten and closely spaced. Subjects were encouraged to look at the digits to be copied rather than at their handwriting so that they would not lose their place in the list.

On each trial the target line was placed at one of three distances from the subject. It was necessary to provide more than one target position to prevent subjects learning the position of the target. In this way subjects were forced to attend to the target on every trial. However, the positions of the target were closely grouped in order to avoid any differential distortion of their perceived positions. The distance between the two extreme positions was 0·8 in. Each subject made six settings of the stylus under each perceptual condition, two for each of the three target positions. The starting position of the stylus was alternated on successive trials between its extreme far and near limits of travel. Two or three practice trials were given during the first session to familiarize subjects with the task. Subjects were instructed to align the stylus with the target line as accurately as they could and were allowed unlimited time for each alignment. They were encouraged to correct their positioning of the stylus if they wished, but no knowledge of results was given.

Results

Inaccuracy of alignment of the stylus with the target could have two kinds of result: a change in the variability of performance, and a systematic constant error. The variability of performance was assessed by calculating the *log variance* (to base 10) of the six settings made under each experimental treatment, the log transformation being used to normalize the distribution. The accuracy of performance was assessed by calculating the *mean error* of the six settings. These two measures were analysed separately.

Variability

The variability of alignment under the sixteen experimental treatments is illustrated in Fig. 1. Each point on the dose-response curves represents the mean of the scores from eight subjects. Mean scores are shown in Table 1. The data were analysed in a three-way classification analysis of variance (Table 2); a mixed model was assumed (case XIV, McNemar, 1955).

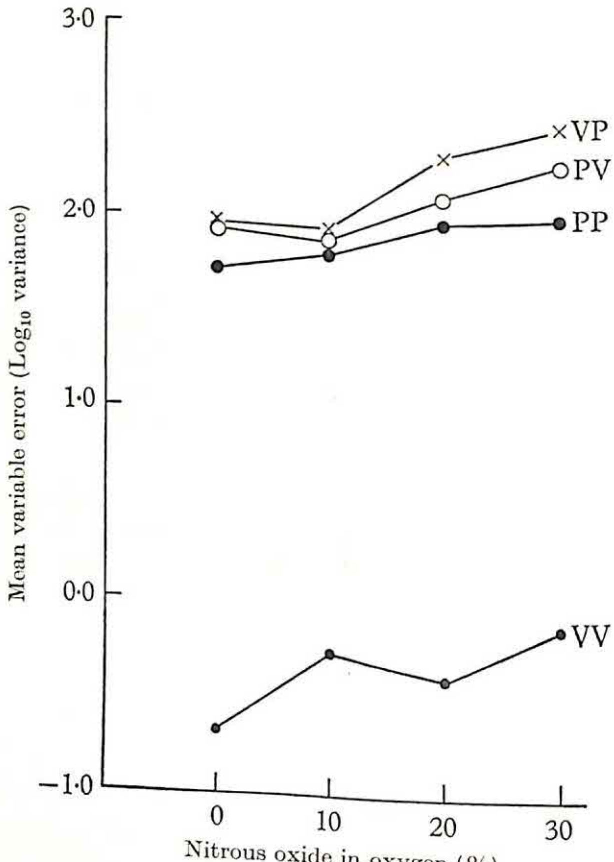


Fig. 1. Variability of alignment for four different perceptual conditions as a function of nitrous oxide dose. Each dose/response curve was obtained under a different perceptual condition. Each point is the mean from eight subjects. VV, both target and stylus perceived visually; VP, target perceived visually, stylus proprioceptively; PV, target perceived proprioceptively, stylus visually; PP, both target and stylus perceived proprioceptively.

Table 1. *Mean variable errors (log_{10} variance, mm.) as a function of dose of nitrous oxide in oxygen for four different perceptual conditions:* VV, *both target and stylus perceived visually;* VP, *target perceived visually, stylus proprioceptively;* PV, *target perceived proprioceptively, stylus visually;* PP, *both target and stylus perceived proprioceptively*

Perceptual conditions	Nitrous oxide in oxygen (%) 0 (air)	10	20	30
VV	−0·675	−0·278	−0·390	−0·128
VP	1·933	1·908	2·298	2·441
PV	1·925	1·858	2·069	1·250
PP	1·708	1·788	1·942	1·971

Three main points emerge. First, the drug clearly produced a general increase in variability of performance. This effect is linear; the larger the dose the greater the variability. Secondly, the different perceptual conditions were characterized by different degrees of variability. When both target and stylus were perceived visually, alignment performance was considerably less variable than when they were both perceived proprioceptively. However, when information from the two modalities had to be combined (conditions VP and PV) performance was more variable still.

Table 2. *Analysis of variance of variability of alignment* (log_{10} *variance*)

Source	SS		D.F.	MS	F	P
Dose	3·140		3	1·047*	6·61	< 0·01
Linear		3·106	1	3·106*	19·65	< 0·001
Quadratic		0·026	1	0·026*	< 1·0	N.S.
Cubic		0·008	1	0·008*	< 1·0	N.S.
Perceptual conditions	136·810		3	45·603†	185·84	< 0·001
VV *vs.* (PP, VP, PV)		135·421	1	135·421†	551·87	< 0·001
PP *vs.* (VP, PV)		1·159	1	1·159†	4·72	< 0·05
VP *vs.* PV		0·230	1	0·230†	< 1·0	N.S.
Subjects	1·647		7	0·239	—	—
Dose × conditions	0·943		9	0·105	1·13	N.S.
Dose × subjects	3·325		21	0·158	1·72	< 0·05
Conditions × subjects	5·153		21	0·245	2·66	< 0·01
Residual	5·800		63	0·092	—	—
Total	156·819		127	—	—	—

* Tested against dose × subjects mean square. † Tested against conditions × subjects mean square. N.S. ≡ $P > 0·05$.

Table 3. *Mean constant errors (mm.) as a function of dose of nitrous oxide in oxygen for four different perceptual conditions.* (cf. *Table.* 1)

Perceptual condition	Nitrous oxide in oxygen (%) 0 (air)	10	20	30
VV	−1·03	−0·87	−0·83	−0·90
VP	−12·95	−8·22	−1·93	−8·00
PV	18·52	8·22	4·64	1·39
PP	−2·13	−5·80	−0·87	−14·30

Thirdly, there was no interaction between dose of nitrous oxide and the perceptual conditions. The increase in variability caused by the drug was unmodified by the perceptual conditions under which the alignment was performed. This lack of evidence of a differential effect of the drug on proprioception and vision cannot be due to the small size of the experiment. The observed value of the Residual Mean Square in Table 2 is equal to the theoretical value calculated from the sample size (Bartlett & Kendall, 1946), indicating that no increase in precision would be obtained by using more subjects.

Constant error

Analysis of constant errors was complicated by the fact that when both target and stylus were perceived visually, stylus settings were considerably less variable than under the other perceptual conditions. This heterogeneity of variance could not be simply removed by a transformation of the scale. By inspection, it was clear that the

Table 4. *Analysis of variance of constant errors of alignment: alignments made when both target and stylus could be seen were omitted from the analysis*

Source	SS		D.F.	MS	*F*	*P*
Dose	989·8		3	329·9	2·55	N.S.
Perceptual conditions	4843·4		2	2423·1	1·64*	N.S.
(*A*) PP *vs.* (VP,PV)		764·6	1	764·6	< 1·0*	N.S.
(*B*) VP *vs.* PV		4078·7	1	4078·7	2·76*	N.S.
Subjects	20234·2		7	2890·6	—	—
Dose × conditions	1705·1		6	284·1	2·19	N.S.
Linear dose × (*B*)		1159·0	1	1159·0	8·96	< 0·01
Residual		546·0	5	109·2	< 1·0	N.S.
Dose × subjects	3667·4		21	174·6	1·35	N.S.
Conditions × subjects	20679·7		14	1477·1	11·42	< 0·001
Residual	5428·3		42	129·2	—	—
Total	57548·2		95	—	—	—

* Tested against conditions × subjects mean square. N.S. ≡ $P > 0{\cdot}05$.

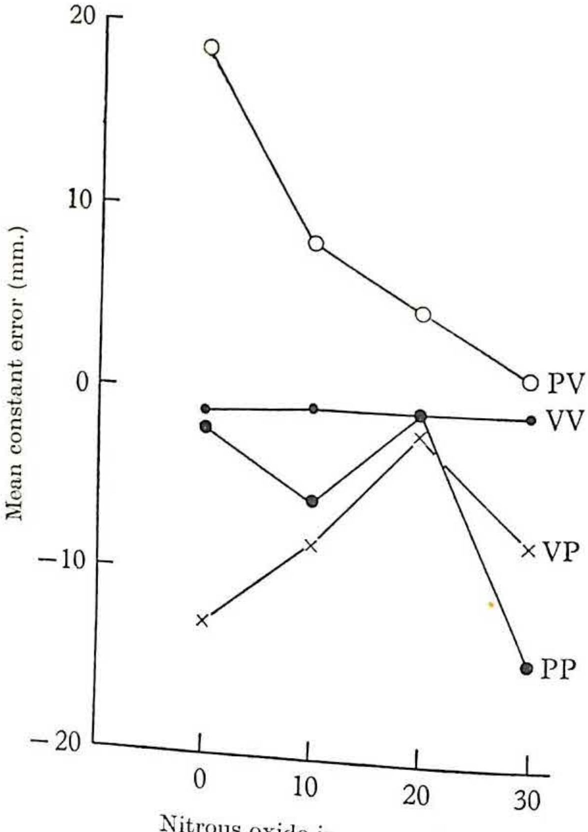


Fig. 2. Constant errors of alignment for four different perceptual conditions as a function of nitrous oxide dose. Key as for Fig. 1.

drug had no effect at all on the stylus settings under this condition (VV), and data collected under the VV condition were omitted from the analysis.

Two main trends in the data are revealed by the analysis of variance (Table 4). First, the difference between constant errors under different perceptual conditions varied from subject to subject. Secondly, nitrous oxide altered differentially the

constant errors under the two cross-modal perceptual conditions. The interaction was a linear one indicating that the underlying drug effect was itself linear.

Fig. 2 shows the effect of three doses of nitrous oxide on the constant errors under the four perceptual conditions. Clearly the drug had no effect when both target and stylus were perceived visually, and it had little effect when they were both perceived proprioceptively. However, the dose-response curves for the two cross-modal perceptual conditions (PV and VP) are very nearly mirror reflexions of each other. When the stylus could be seen but the target could not, the constant errors became

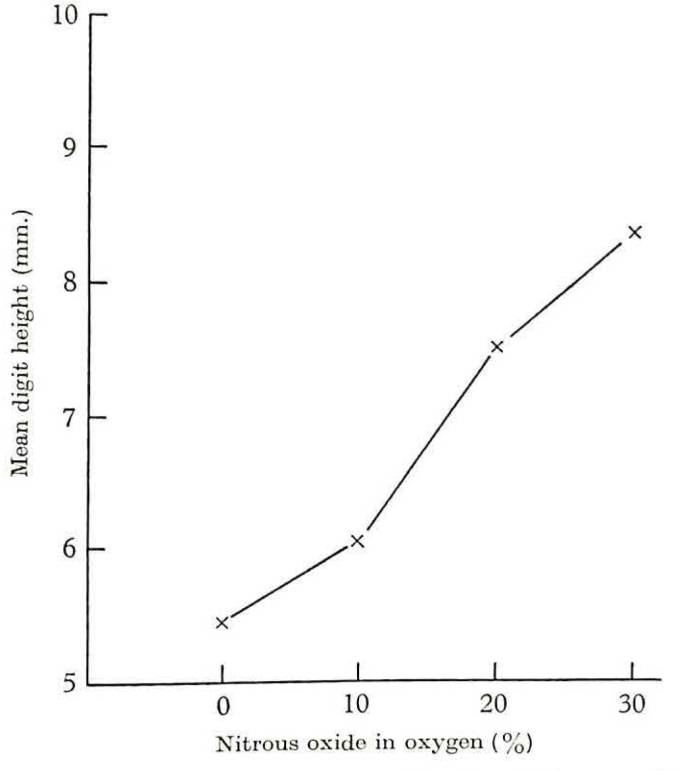


Fig. 3. Height of handwritten digits as a function of nitrous oxide dose.

progressively more negative as the dose increased. When the target could be seen but the stylus could not, the constant errors became more positive with increase in dose. When the stylus could not be seen this simple relation was complicated by an unexpectedly large negative constant error at the highest dose. This unexpected negative constant error also occurred when neither the target nor the stylus could be seen. It seems likely that these anomalous errors under conditions VP and PP may be attributed to the way in which the stylus was perceived proprioceptively at the highest dose of nitrous oxide. The stylus had to be both located in space and moved, while the target had only to be perceived. The highest dose may have markedly reduced tactual sensitivity and it is possible that subjects gripped the stylus in an unusual way in order to move it. This change of hand-grip may have altered the perceived position of the stylus and hence produced these unexpected constant error scores.

Handwriting

Fig. 3 shows the effect of nitrous oxide on the height of handwriting. The vertical height was measured of the last five figure 'nines' in the list of random numbers which was copied by the subject at the beginning of the experiment. This measure was shown to be a relatively sensitive index of drug effect in an analysis of the effects of nitrous oxide on handwriting by Legge *et al.* (1964). Handwriting size increased as a linear function of dose of nitrous oxide ($F = 29{\cdot}44$; D.F. 1, 21; $P < 0{\cdot}001$).

DISCUSSION

The results of the experiment show that the variability with which subjects aligned the stylus with the target was affected by the perceptual conditions imposed on the subject. Variability of performance was also affected by nitrous oxide. Variability increased as a linear function of the dose of nitrous oxide. There was no interaction between the perceptual conditions and the dose of drug, which indicates that the effects of the drug on variability were independent of perceptual modality. It appears that the drug acted upon some process common to both modalities. The drug may have acted on processes subserving judgement and increased subjects' tolerance of errors of alignment. Under the influence of the drug subjects accepted misalignments which they corrected when performing without the drug.

The effects of different combinations of perceptual modalities on variability of alignment were more complex. Vision is a more acute modality than proprioception and it is no surprise that when both target and stylus were perceived visually, performance was more precise than under any of the conditions involving proprioception. However, it is surprising that when vision and proprioception had to be combined performance was more variable than when both target and stylus were perceived by the less acute modality, proprioception. The implication is that when performance depends upon the integration of visual and proprioceptive spatial information some additional process is introduced. This additional process is imperfect and its inclusion increases the variability of performance.

A combination of proprioceptive and visual cues is generally thought to be the normal basis of the control of motor skills of many kinds (Fitts, 1951; Pieron, 1952; Hefferline, 1958). It is therefore surprising that these two modalities should be less efficient in combination than either is alone. This problem has received little attention. Rudel & Tueber (1963) report that the decrement in the Muller–Lyer illusion in man which may be induced simply by exposure of the subject to the illusion, transfers from touch to vision and vice versa. Ettlinger (1960), however, failed to discover any transfer of a shape discrimination between these modalities in rhesus monkeys. It may be that the process underlying the decrement in the amount of the Muller–Lyer illusion is a relatively high level one, already far removed from the primary representation of information in either modality. Alternatively, the critical difference between these two experiments may have been the species of the experimental subjects. Kelvin & Mulik (1958) used the constant method to study cross-modal perception of the size of small disks. The results suggest that subjects found the cross-modal comparison so difficult that they ignored the standard entirely and made their responses simply in terms of the set of comparison stimuli.

There have been few studies of the precision of cross-modal localization, but G. H. Fisher (personal communication) and Merton (1961) both report greater variability when visual information is compared with proprioceptive information in a direction-indicating task than when similar comparisons are made within either modality. These findings are consistent with the results reported here.

The constant error data from this experiment illuminate the relation between visual and proprioceptive representations of the same physical location. For the subjects in this experiment there was a consistent mismatch between visual and proprioceptive spatial localizations. This mismatch is illustrated in Fig. 2 in which constant error scores are shown for the two cross-modal perceptual conditions. In general the visually perceived stylus had to be placed closer to the subject than did the proprioceptively perceived target in order that the stylus should be perceived in line with the target. The opposite relation held when the opposite combination of modalities obtained. A similar result was reported by Sandstrom (1951) although his particular sample of subjects produced a distribution of constant errors with a mean of opposite sign. Systematic differences between visual and proprioceptive impressions of size are reported by a number of authors including Jastrow (1886), Abel (1936) and Raffel (1936). Variation in the sign and the size of the mean constant errors among different groups of subjects shows that although the cross-modal relation for a particular subject may be stable, there are large individual differences in this relation.

The evidence, both past and present, strongly suggests that there is no simple way in which information mediated by vision and proprioception may be integrated. First, there is increased variability when these modalities have to be involved in a single performance which suggests the introduction of an additional mediating process. This process might translate information from the terms and units of one modality into the terms and units of the other. Secondly, there is good evidence of a basic mismatch between these modalities. The role of proprioception in motor skills clearly needs to be re-examined.

Fig. 2 shows that as the dose of nitrous oxide was increased the visual-proprioceptive constant error altered systematically. There are two simple hypotheses which could account for this result. The drug may have acted to increase visual apparent distance, or to decrease proprioceptive apparent distance, or to produce both effects simultaneously. It cannot be inferred from this experiment which modality was affected by the drug.

Steinberg, Legge & Summerfield (1961) have shown that nitrous oxide alters judgement of size at a distance; a result which might suggest that the drug affects visual apparent distance. A near variable stimulus was to be set to appear equal in size to a far standard. The subject manipulated the near variable stimulus with his right hand. The effect of giving the drug could have been to produce a disproportionate increase in the apparent distance of the far standard stimulus. However, while the distance of the far stimulus was perceived visually, that of the near variable stimulus was perceived both visually and proprioceptively. Thus it is possible that the drug-induced increase in the adjusted size of the near stimulus could have been due to its being perceived as closer than it really was. Alternatively, the drug could have acted upon the central mechanisms subserving constancy rather than upon the perception

of distance. The problem of determining which modality was affected by nitrous oxide remains unsolved.

Although the effects of nitrous oxide on perception are not yet fully understood it is instructive to examine how far the drug-induced increase in handwriting size may be explained in terms of the changes in perception observed in the experiment reported here. Three effects on perception may be considered; an increase in visually perceived size possibly caused by an increase in visual apparent distance, a decrease in proprioceptive apparent distance, and a decrease in the precision of information mediated by either modality. Each effect could produce an increase in handwriting size.

Handwriting samples were obtained from subjects who copied either a passage of prose (Legge *et al.* 1964) or typewritten digits (*vide supra*). If size of handwriting is a function of the apparent size of the typed sample an increase in the apparent size of the letters or digits to be copied would produce an increase in handwriting size. The weakness of this explanation is that it depends on the assumption that the size of subjects' handwriting is influenced by the apparent size of the material they are copying.

A more plausible explanation may be advanced in terms of a drug-induced decrease in the proprioceptively perceived size of handwriting. If handwriting is normally maintained at a particular size by some monitoring system, an increase in handwriting size would result if the monitoring system 'reported' handwriting to be too small. Subjects were encouraged to look at the material to be copied rather than at their handwriting and proprioception would have to have been used if handwriting were monitored continuously. Hence the observed increase in handwriting size could have resulted from a decrease in apparent handwriting size induced by the action of the drug on proprioception.

A third explanation may be advanced in terms of the ease with which handwritten characters may be identified proprioceptively. It has been shown that nitrous oxide increases the variability of perceptual performance. This effect would reduce proprioceptive acuity. If subjects monitor their handwriting proprioceptively, one feature that it would be important to monitor would be the nature of each character written. The proprioceptive legibility of a written character will be a function of its size and any decrease in proprioceptive acuity would reduce the ease with which a character of a particular size could be identified. Thus the observed increase in handwriting size could result from compensation for a drug-induced reduction in proprioceptive acuity.

Clearly further experiment is necessary to determine which of these explanations should be accepted. Experiments are also necessary to determine how far the effect of these drugs on handwriting size can be attributed simply to effects on perception. Other features of motor skill may be involved such as the speed and accuracy of stimulus registration and the retrieval of stored information (Summerfield, 1964), the timing of complex sequences of movements (Bartlett, 1948; Provins, 1956), and the accuracy with which a central representation of the response is prepared (Poulton, 1957; Begbie, 1959; Henry & Harrison, 1961). It would also advance the problem to discover whether a single explanation could be generalized from nitrous oxide to other depressant drugs like alcohol which also increase handwriting size (Rabin & Blair, 1953; Tripp, Fluckiger & Weinberg, 1959).

I am most grateful to Prof. A. Summerfield and Dr H. Steinberg for their encouragement, advice and practical direction, to Prof. D. R. Cox for statistical advice, to Dr M. Lader for advice and assistance on medical issues, to my colleagues for many helpful discussions and to the subjects who so patiently gave their time.

This investigation was supported in part by research grant MH-03313 from the National Institute of Mental Health, Public Health Service, U.S.A.

References

Abel, T. M. (1936). A comparison of tactual-kinaesthetic and visual perceptions of extent among adults, children and sub-normals. *Amer. J. Psychol.* **48**, 269–96.

Bartlett, F. C. (1948). The measurement of human skill. *Occ. Psychol.* **22**, 31–8.

Bartlett, M. S. & Kendall, D. G. (1946). The statistical analysis of variance—heterogeneity and the logarithmic transformation. *J. R. statist. Soc.* B, **8**, 128–38.

Begbie, G. H. (1959). Accuracy of aiming in linear hand-movement. *Quart. J. exp. Psychol.* **11**, 65–75.

Carpenter, J. A. (1962). Effects of alcohol on some psychological processes. *Quart. J. stud. Alc.* **23**, 274–413.

Drew, G. C., Colquhoun, W. P. & Long, H. A. (1959). *Effect of Small Doses of Alcohol on a Skill Resembling Driving*. London: H.M.S.O.

Ettlinger, G. (1960). Cross-modal transfer of training in monkeys. *Behaviour*, **16**, 56–65.

Fitts, P. M. (1951). Engineering psychology and equipment design. In *Handbook of Experimental Psychology*, ed. Stevens, S. S. New York: Wiley.

Fleishman, E. A. & Rich, S. (1963). Role of kinaesthetic and spatial-visual abilities in perceptual-motor learning. *J. exp. Psychol.* **66**, 6–11.

Gibbs, C. B. (1954). The continuous regulation of skilled responses by kinaesthetic feedback. *Brit. J. Psychol.* **45**, 24–39.

Goodman, L. S. & Gilman, A. (1955). *The Pharmacological Basis of Therapeutics*. New York: Macmillan.

Hefferline, R. F. (1958). The role of proprioception in the control of behaviour. *Trans. N.Y. Acad. Sci.* **20** Ser. 2, 739–64.

Henrie, J. H., Duncan, G. H. & Rome, H. P. (1960). N_2O analgesia in relation to mental performance. *J. Pharmacol.* **128**, 44–54.

Henry, F. H. & Harrison, J. S. (1961). Refractoriness of a fast movement. *Percept. Mot. Skills*, **13**, 351–4.

Jastrow, J. (1886). The perception of space by disparate senses. *Mind*, **11**, 539–54.

Jellinek, E. M. & McFarland, R. A. (1940). Analysis of psychological experiments on the effects of alcohol. *Quart. J. Stud. Alc.* **1**, 272–371.

Kelvin, R. P. & Mulik, A. (1958). Discrimination of length by sight and touch. *Quart. J. exp. Psychol.* **10**, 187–92.

Kety, S. S. & Schmidt, C. F. (1948). The nitrous oxide method for the quantitative determination of blood flow in man. *J. clin. Invest.* **27**, 476–83.

Kornetsky, C., Humphries, O. & Evarts, E. V. (1957). Comparison of psychological effects of certain centrally acting drugs in man. *Arch. Neurol. Psychiat.* **77**, 318–24.

Legge, D., Steinberg, H. & Summerfield, A. (1964). Simple measures of handwriting as indices of drug effects. *Percept. Mot. Skills*, **18**, 549–58.

McNemar, Q. (1955). *Psychological Statistics*. New York: Wiley

Merton, P. A. (1961). The accuracy of directing the eyes and the hand in the dark. *J. Physiol.* **156**, 555–77.

Payne, R. B. & Moore, E. W. (1955). The effects of some analeptic and depressant drugs upon tracking behaviour. *J. Pharmacol.* **115**, 480–4.

Piéron, H. (1952). *The Sensations*. New Haven: Yale University Press.

Poulton, E. C. (1957). On prediction in skilled movements. *Psychol. Bull.* **54**, 467–78.

Provins, K. A. (1956). Handedness and skill. *Quart. J. exp. Psychol.* **8**, 79–95.

Rabin, A. & Blair, H. (1953). Effects of alcohol on handwriting. *J. clin. Psychol.* **9**, 284–7.

Raffel, G. (1936). Visual and kinaesthetic judgements of length. *Amer. J. Psychol.* **48**, 331–4.

Rudel, R. G. & Teuber, H. L. (1963). Decrement of visual and haptic Muller–Lyer illusion on repeated trials: a study of cross-modal transfer. *Quart. J. exp. Psychol.* **15**, 125–31.

SANDSTROM, C. I. (1951). *Orientation in the Present Space*. Stockholm: Almquist, Wiksell.

STEINBERG, H. (1954). Selective effects of an anaesthetic drug on cognitive behaviour. *Quart. J. exp. Psychol.* **6**, 170–80.

STEINBERG, H., LEGGE, D. & SUMMERFIELD, A. (1961). Drug-induced changes in visual perception. In *Neuro-Psychopharmacology*, vol. 2, pp. 392–6., ed Rothlin, E. Amsterdam: Elsevier.

SUMMERFIELD, A. (1964). Drugs and human behaviour. *Brit. med. Bull.* **20**, no. 1, 70–4.

TRIPP, C. A., FLUCKIGER, F. A. & WEINBERG, G. H. (1959). Effects of alcohol on the graphomotor performances of normals and chronic alcoholics. *Percept. Mot. Skills*, **9**, 227–36.

WILLIAMS, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian J. sci. Res.* A, **2**, 149–68.

(*Manuscript received* 1 *April* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp*. 255–265
Printed in Great Britain

EXPERIMENTAL ANALYSIS OF DRUG EFFECTS ON HUMAN PERFORMANCE USING INFORMATION THEORY CONCEPTS

By COLIN BERRY, MICHAEL G. GELDER and ARTHUR SUMMERFIELD*

Birkbeck College and University College London

Effects of differences in mean stimulus information under two coding conditions and of sub-anaesthetic doses of nitrous oxide (15, 25 and 35% in oxygen) were investigated in two card-sorting experiments with student subjects In Expt. I, in which conventional playing cards were sorted into two, four or eight classes, the effect of the drug increased significantly with task complexity. Expt. II, in which cards bearing numerals were used, showed a drug effect which was independent of task complexity measured by mean information per stimulus. Neither result was to be explained in terms of a drug effect on the motor component of the tasks. Reasons for the difference between the two experiments are considered in relation to other evidence of effects of central nervous depressant drugs on input processes and short-term memory. The value of communication models for research on effects of drugs on human skills is discussed.

Differential effects of drugs on human behaviour can facilitate analysis of the behaviour into component processes. A characteristic of more complex forms of behaviour is that they involve sequences of receptor and effector processes organized in time, rather than isolated reactions. Information measures provide a convenient basis for comparing task complexity and efficiency of performance for sequential tasks (Summerfield & Legge, 1960, 1964), and for expressing dose–response relations in analysing drug effects. Concepts derived from information and communication theory facilitate analysis of drug effects in other ways (Berry, Gelder & Summerfield, 1963). Processes like perception, discrimination, choice and motor co-ordination can be discussed in common terms of 'coding' inputs into outputs. Coding in this sense can be broken down into a sequence of stages, or mechanisms, intervening between input and output. Schemes of this kind have been suggested by Welford (1960) and, more elaborately, by Crossman (1964). Perception, discrimination, choice and so on are interrelated by this approach, instead of tending to be discussed as distinct processes. Furthermore, experiments can be devised with the object of looking for selective effects on different stages in the functional chain by different drugs. This plan involves using tasks with different degrees of dependence on the several stages. Two kinds of task have been used, sorting tasks (cf. Crossman, 1953) which are discussed here, and tasks involving simple numerical coding (cf. Alluisi & Martin, 1958; Summerfield, 1964*a*, *b*).

Steinberg (1959) has suggested that central depressant drugs may act selectively: (*a*) to impair newly acquired rather than well-established behaviour, (*b*) to affect different kinds of performance differently, for example motor tasks and more intellectual tasks, and (*c*) to have a greater effect on performance, the more complex the task. For some kinds of task, including sorting tasks, complexity can be measured

* Dr M. G. Gelder, who was Medical Research Council Fellow in Clinical Research at Birkbeck College, is now at the Institute of Psychiatry, University of London. The work was supported in part by research grant MH-03313 from the National Institute of Mental Health, Public Health Service, U.S.A.

in terms of the 'average information per signal', or per stimulus, in a number of stimulus classes. An advantage is that the average information per stimulus, in *bits*, which is associated with each level of task complexity, tends to be related linearly to response time when time is used as the measure of performance. This relation is a convenient base-line against which to evaluate drug effects.

The first experiment explored the usefulness of a card-sorting task for investigating effects of the central nervous depressant drug nitrous oxide. The task (Crossman, 1953) required subjects to sort a pack of playing cards into different piles according to the symbols on each card. Although Crossman originally used conventional playing cards, other kinds of visual material can be used. A task of this kind involves not only selective coding in the sense of choice or decision, but also the movement of placing the cards on appropriate piles. While for many purposes a large movement factor is not desirable, a sorting task allows comparison of drug effects on the two elements of choice and movement. Crossman mentions another difficulty which arises when conventional playing cards are used; it is not possible to vary information without also changing other factors which may have an effect on performance. For this reason, after a preliminary experiment with conventional cards, different stimuli were used in a second experiment.

EXPERIMENT I

Method

Plan. There were four sorting conditions in this experiment. Each pack to be sorted was of thirty-two conventional playing cards. One pack contained sixteen red and sixteen black number cards and was sorted by subjects into two piles according to colour alone, other features being ignored. The second pack contained eight cards of each suit and was sorted into four piles. The third pack was made up in the same way as the second, the four suits being represented by four number and four court cards; this pack was sorted into eight piles, being divided into court and number cards as well as into suits. The 2-, 4- and 8-choice tasks required the use of 1, 3 and 3 bits of information/stimulus card. In a fourth condition (M), subjects dealt the cards by turning them over and placing them alternately on two piles, regardless of the symbols on the cards, so that the movement element in the 2-choice task could be investigated. This procedure was repeated with some subjects for 8 choices, but not for 4. The measure of performance was the time taken to sort all the cards in each case.

Nitrous oxide was used because it is easily given, is safe, and produces reliable effects which are induced and disappear rapidly without after effects, so that different doses can be given on successive days. It is, therefore, particularly suitable for investigations using new techniques. Twenty-five per cent nitrous oxide in oxygen was considered likely to produce measurable effects without causing distress or loss of co-operation in the majority of subjects. Air was used as a control. Subjects breathed the gases, scented with lavender oil, through a face mask. The technique first described by Steinberg (1954) was used.

The order of the tasks was the same for all subjects. A simple reversed series which was intended to balance fatigue and practice effects was adopted: M, 4, 2, 8, 8, 2, 4, M. Subjects did the whole series twice. For the first series all subjects breathed air; for the second series experimental subjects were given the drug and control subjects again had air. A standard pause of 5 min between the two halves of the session allowed for equilibriation to the drug by the experimental group.

Subjects. Eighteen students (eight men and ten women, mean age 20·7 yr, range 18–30 yr) were randomly assigned to two equal groups, an experimental and a control group. Three subjects had had previous experience of nitrous oxide in subanaesthetic doses in psychological experiments; these three fell by chance into the control group. Standard questions were asked about medical history. No subject who was under treatment with other drugs was admitted to the experiment.

Procedure. The subject sat at a table holding the well-shuffled pack of cards face down, and was instructed to turn them over one at a time and to place them as quickly as possible in piles on the table. Marker cards were not used and subjects were free to place the piles where they chose. They first practised the three sorting conditions once, in ascending order of difficulty, while breathing air through the face mask. They were encouraged to perform as rapidly as was consistent with accuracy, to avoid errors as far as possible, but not to stop or make corrections. No more than two errors per pack were allowed. (In spite of the instructions, subjects generally made corrections which were not scored; only one run had to be repeated out of 144.) Subjects were told that they would be breathing scented air in the first half of the main experiment and that in the second half they would either continue to have air or would be given a low dose of nitrous oxide. They were told that they might perhaps feel a little strange but that they would not lose consciousness and should persevere with the task. Although subjects were not paid they seemed to be highly motivated; they were, for example, eager to do well and to discuss their results after the experiment.

Results

In the first period, when both groups breathed air, the mean sorting times for the control group were slightly longer than those for the experimental group. But this relation was reversed in the second period when the experimental group had the drug. As average times for the two groups were different in the first half of the experiment, difference scores were used to evaluate the drug effects: the two times for each subject on each task in each half of the experiment were first averaged; averages for the first half were subtracted from the corresponding averages for the second half of the experiment.

Fig. 1 shows these differences for the two groups under the three sorting conditions. Differences for the control group are negative, indicating that subjects given air sorted more quickly on the second occasion: differences increased little with increasing stimulus information implying that the effects of practice between the two periods, which are small, are almost the same for the three levels of difficulty. Differences for the experimental group, however, are positive: sorting times were longer on the second occasion and the differences increased with the number of stimulus classes.

Mean values for the two conditions are given in Table 1. Analysis of variance shows significant differences between drug and air conditions, and between stimulus information conditions. The drug × stimulus information interaction is also significant, indicating that the effect of the drug differs at the three information levels. Regression equations confirm that the difference scores for the air group do not increase significantly with increasing stimulus information, but show that there is a significant linear regression for the drug group ($F = 6{\cdot}59$, D.F. 1, 24; $P < 0{\cdot}05$; no significant departure from linear regression). The test of linearity does not reach a high level of significance because the mean times under air for the experimental group themselves depart from linearity (cf. Table 1); the mean times under drug are more linear. The mean values in Table 1 show that the apparent discrepancy in the difference scores results from inconsistency in the means of the data for the experimental group under the initial air condition.

Fig. 1 shows that 25 % nitrous oxide has little effect on movement in this task, suggesting that the increases in sorting times may result from an effect of the drug on some central coding process rather than on the motor responses.

More rigorous repetition of this experiment would be desirable if it were not for

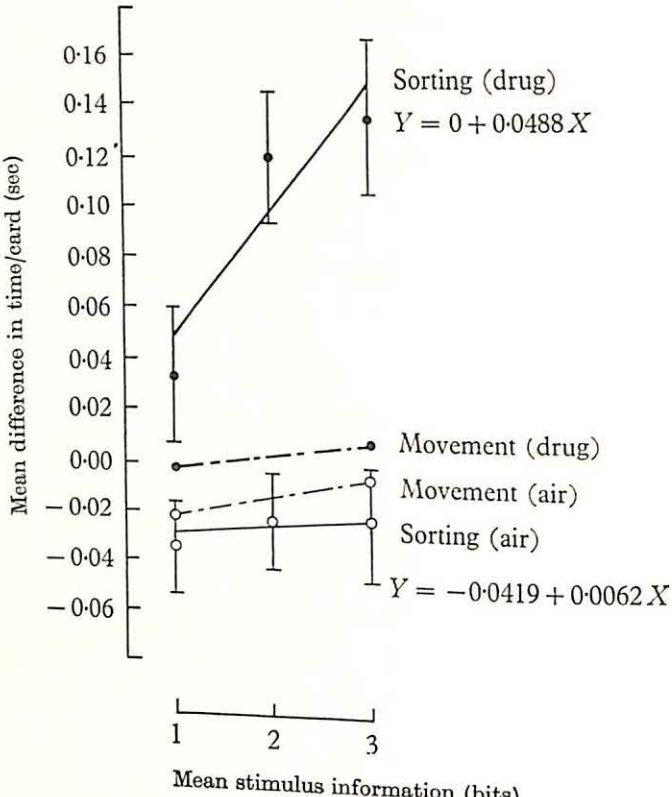


Fig. 1. Regressions of mean differences in sorting times/card (32 cards) on stimulus information for three sorting tasks, with mean differences in times for movement only for comparison, by the drug and air (control) groups of Expt. I ($n = 9$ per group). The drug group breathed 25 % nitrous oxide in oxygen during the second half of the experiment.

Table 1

A. *Mean times (sec) to sort 32 cards under three sorting conditions (1, 2, 3 bits/stimulus) and a control movement condition (M): subjects in the drug group breathed 25 % nitrous oxide in oxygen during the second half of the experiment*

Mean stimulus information	Air group ($n = 9$)			Drug group ($n = 9$)		
	1st half (air)	2nd half (air)	Difference	1st half (air)	2nd half (drug)	Difference
1 bit	28·10	26·89	−1·21	25·66	26·81	1·15
2 bits	35·49	34·66	−0·83	34·20	38·07	3·87
3 bits	46·54	45·76	−0·78	45·40	49·85	4·45
M	(16·08)	(15·21)	(−0·87)	(15·20)	(15·22)	(0·02)

B. *Analysis of variance for the three sorting conditions*

Source of variation	S.S.	D.F.	M.S.	F	P
Drug condition	97·54	1	97·54	25·30	< 0·001
Error (a)	61·76	16	3·86	—	—
Stimulus information	38·00	2	19·00	3·03	< 0·05
S × D	144·43	2	72·21	11·51	< 0·001
Error (b)	200·50	32	6·27	—	—
Total	542·23	53	—	—	—

doubt about the stimulus material. The subject has to attend to different aspects of the cards in the three tasks. Mean stimulus information, as we have defined it, therefore depends on several differences. For example, some of the information on the cards which is used in sorting into eight classes becomes irrelevant when the same cards are sorted in four classes. There is a difference in coding between deciding whether a card is red or black and deciding whether it is a court or a number card; and it is possible that the drug might affect such perceptual discriminations. A second experiment was therefore planned which was free from these complications.

Experiment II

Method

Plan. Special packs of cards were prepared for this experiment: blank playing cards were used and arabic numerals between '1' and '8', 1 in. high, were stencilled on them. There were forty-eight cards in each of four packs which were made up as follows: (i) twenty-four each of '1' and '2', (ii) twelve each of '1', '2', '3' and '4', (iii) eight each of '1' to '6', and (iv) six each of '1' to '8'. Numerals were chosen because they enabled each card to carry a single symbol of one kind, i.e. digits, unlike the more complex playing cards. They are also familiar symbols, giving a stereotyped relation between the stimulus cards and the ordered sequence into which they can be sorted, i.e. a high degree of S–R compatibility (Alluisi & Muller, 1958; Deininger & Fitts, 1955). In addition, they enable comparisons to be made with other experiments in which numerals are used as stimuli (cf. Summerfield, 1964*a*, *b*; Posner, 1965).

There were therefore four task conditions. The four packs were used respectively for sorting or dealing into 2, 4, 6 and 8 classes, so that the mean stimulus information per card for sorting the well shuffled packs was 1, 2, 2·59 or 3 bits. In order again to investigate movement all subjects were required to deal the cards into piles, as well as to sort them. The measure of performance was the time taken to sort or deal each pack of forty-eight cards. Two kinds of errors were also recorded: misplacements and dropped cards.

Table 2. *Plan of Expt. II: Williams's square design (cf. Cox, 1961) for four tasks* $(a, b, c, d) \times$ *four doses* (D_1, D_2, D_3, D_4) *for eight subjects on 4 days*

Days ...	I	II	III	IV	
Periods ... / Subjects	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	
1	D_1 *a b c d*	D_4 *d c b a*	D_3 *c a d b*	D_2 *b d a c*	I
2	D_2 *d c b a*	D_3 *a b c d*	D_4 *b d a c*	D_1 *c a d b*	I
3	D_3 *b d a c*	D_1 *d c b a*	D_2 *c a d b*	D_4 *a b c d*	I
4	D_4 *c a d b*	D_2 *a b c d*	D_1 *b d a c*	D_3 *d c b a*	I
5	D_1 *c a d b*	D_4 *b d a c*	D_3 *d c b a*	D_2 *a b c d*	II
6	D_2 *b d a c*	D_3 *c a d b*	D_4 *a b c d*	D_1 *d c b a*	II
7	D_3 *a b c d*	D_1 *b d a c*	D_2 *d c b a*	D_4 *c a d b*	II
8	D_4 *d c b a*	D_2 *c a d b*	D_1 *a b c d*	D_3 *b d a c*	II

Three doses of nitrous oxide were used: 35, 25 and 15% in oxygen, with air as a control. Previous experience suggested that 35% was about the largest dose that could be given without losing the co-operation of some subjects; 25% allowed comparison with Expt. I. The gases were given as before. Subjects were given a different dose on each of 4 consecutive days.

The combination of four tasks, four doses and four days allowed the use of Williams's square (cf. Cochran & Cox, 1957), as modified by Legge (cf. Cox, 1961; Legge, 1965). Williams's square has the property that each condition occurs an equal number of times in each position, as in the Latin square; in addition, each condition immediately follows every other condition an equal number of times. This eliminates order effects and has the advantage over the Latin square that it allows the residual effects of each treatment to be estimated, and corrections to be applied if necessary. The plan which was used is given in Table 2. The four dose conditions are ordered in superordinate Williams's squares I and II. The four task conditions are ordered in eight further squares, two for each dose condition, which are incorporated in the overall plan. In each task condition the subject first dealt and then sorted the pack. The four main columns of Table 2 show the treatments on successive days; each daily session is divided into four periods. Each row therefore shows the sequence of treatments within periods for one subject. Subjects were allocated at random to the rows of the plan, and as a result it was possible to estimate residual effects of both dose and task conditions.

Subjects. Eight students took part (five men and three women, mean age 20·5 yr, range 18–25 yr). All had taken part in previous experiments on card-sorting under nitrous oxide. Their expenses were paid.

Procedure. As before, subjects sat at a table and turned the cards over one by one. In this experiment marker cards were laid out in numerical order from left to right on the table, and subjects were instructed to place their cards on top of them. The instructions again emphasized speed and accuracy, and discouraged correction of errors. A maximum of three errors per pack was allowed (only one subject made this number). The experimenter began each run by calling 'Ready go!' and starting a stop-watch. When dealing, subjects were encouraged to sit back, holding the pack in one position, and to make a distinct movement for each card rather than to move the pack along the row of cards. This reproduced more closely the movements made when the cards were actually sorted. The first session included a standard practice period in which subjects breathed through the face mask and dealt and sorted each pack once in the order 2, 4, 6, 8.

Results

Performance times were subjected to logarithmic transformation to normalize their distributions. Preliminary analyses showed that the residual effects between days were not significant. Separate analyses for each dose, combining results for the 4 days, showed that the residual effects between Tasks also were not significant.

The analysis of variance on *sorting times* (Table 3*B*) show highly significant effects of tasks, i.e. stimulus information, doses and subjects. The tasks × doses interaction is not significant. The significant doses × subjects and tasks × subjects interactions confirm the expectations of individual differences in response to the drug and in performance respectively.

An identical analysis of *movement times* showed no significant effects of tasks ($F < 1$), or of doses ($F = 1{\cdot}71$, D.F. 3, 63; $P > 0{\cdot}10$), but significant differences between subjects ($F = 15{\cdot}25$, D.F. 7, 63; $P < 0{\cdot}001$). The tasks × doses interaction again was not significant. A significant doses × subjects interaction ($F = 2{\cdot}31$, D.F. 21, 63; $P < 0{\cdot}025$) revealed individual differences in susceptibility to the drug, as with sorting times. Thus, while differences in stimulus information between the sorting tasks had a significant effect on sorting time, the corresponding differences between the movement tasks had no effect on movement times. Fig. 2 shows that sorting times increased progressively with increasing dose, but that movement times did not.

Regression analyses for sorting time on stimulus information at each dose showed no significant departures from linearity; moreover, there were no significant differences between the regression coefficients and analysis of variance applied to the regression for each dose revealed no differences in linear regression. The four regres-

Table 3. *Experiment II*

A. *Mean times (sec) to sort 48 cards under four sorting conditions (1, 2, 2·59, 3 bits/stimulus) and four movement conditions at four dose levels of a drug*

	No. of classes	Mean stimulus information bits	Dose (% nitrous oxide in oxygen) 0% (air)	15%	25%	35%
Sorting	2	1·00	40·7	41·9	42·9	47·2
	4	2·00	46·8	48·8	50·8	53·9
	6	2·59	51·5	54·3	55·6	58·5
	8	3·00	55·3	57·5	60·1	63·6
Movement	2	—	24·2	24·3	22·9	24·4
	4	—	22·1	22·9	23·2	28·7
	6	—	22·2	22·8	22·3	25·3
	8	—	22·2	25·0	22·8	25·9

B. *Analysis of variance for the four sorting conditions (T) on transformed sorting times (100 log time)*

Source of variation	S.S.	D.F.	M.S.	F	(Error)	P
Sorting tasks (T)	3420·31	3	1140·10	65·7	(*a*)	< 0·001
Doses (D)	522·89	3	194·30	17·6	(*b*)	< 0·001
Subjects (S)	1369·57	7	195·65	—	—	—
Interactions						
T × D	8·80	9	0·98	< 1	(*c*)	N.S.
T × S (*a*)	364·42	21	17·35	5·56	(*c*)	< 0·001
D × S (*b*)	231·47	21	11·02	3·53	(*c*)	< 0·001
Residual (*c*)	196·76	63	31·23	—	—	—
Total	6174·22	127	—	—	—	—

Table 4. *Multiple regression of transformed sorting times (100 log time) on tasks (stimulus information) and dose*

Source of variation	S.S.	D.F.	M.S.	F	P
Regression on					
Tasks	341·0	1	341·0	246·2	< 0·001
Increment from dose	55·5	1	55·5	40·1	< 0·001
(Tasks + dose)	396·5	2	198·4	—	—
Residual	173·4	125	1·4	—	—

sion lines in Fig. 2 are therefore parallel. Dose–response relations were also examined by regression analyses for sorting time on dose. At each level of stimulus information the regression due to pooled slope was significantly different from zero ($P < 0{\cdot}01$), but again differences between slopes were not significant. The dose–response lines were therefore effectively parallel also (cf. Summerfield, 1964*a* fig. 1, p. 73). The multiple regressions of sorting time on both stimulus information and dose showed that a highly significant component of the total variation was attributable to the effect of dose (cf. Table 4). Stimulus information, however, accounted for a larger share of the total variation than did dose.

Errors were few under all conditions (cf. Table 5). In the sorting task there was some tendency for misplacements to increase at the higher doses, but no tendency for

the number of cards dropped to do so. These results are consistent with the other finding and again imply that in this task the drug mainly impairs information processing rather than movement.

Table 5. *Incidence of errors (dropped and misplaced cards) in Expt. II*

Dose ($\%N_2O$ in O_2)	A. Dropped cards: No. of stimulus classes 2	4	6	8	Total	B. Misplacements: No. of stimulus classes 2	4	6	8	Total
0 (Air)	0	0	2	1	3	2	2	0	4	8
15	0	0	1	0	1	0	2	2	3	7
25	0	0	0	1	1	4	2	2	7	15
35	1	0	0	0	1	2	5	4	2	13

DISCUSSION

The results of the second experiment do not support the hypothesis that the effects of a depressant drug increase disproportionately with increasing task complexity, as defined by mean stimulus information: sorting times increased with increasing stimulus information at each dose, and sorting times were also longer the higher the dose (Fig. 2); but the linear regressions of sorting time on stimulus information did not differ, and in particular the coefficients did not increase with dose. There was therefor no interaction between task complexity and dose of a kind which would have implied an effect of the drug on choice. This is the main difference between Expts. I and II.

In Expt. II, movement time increased slightly with dose between the 25 and 35 % doses, but Fig. 2 shows that there was no change over the lower doses. This disjunction between movement and sorting implies that the longer sorting times with increasing dose are not to be explained by effects of the drug on motor function. Sorting times were however impaired by the lowest dose of nitrous oxide (15 %), and the effect was that each dose of the drug produced a constant delay to give the parallel regression lines. It therefore appears that the principal effect is on a pre-motor process other than the selective coding involved in final choice.

If this is so, the interaction of dose and stimulus information in Expt. I needs to be explained. There were differences between the experiments in plan, stimulus materials and procedure. A pilot experiment which used the materials of Expt. II, but which followed the plan and procedure of Expt. I gave results like those of Expt. II; the difference does not therefore seem to be attributable to differences in experimental plan, a conclusion which is reinforced by the absence of residual effects in Expt. II. Differences in the stimulus materials have already been discussed: in Expt. I subjects had to attend to different aspects of the playing cards in the three tasks and more discriminations were needed in the more complex tasks, e.g., in the 8-choice task colours, suits and values all had to be distinguished. The two experiments also differed in the use of marker cards which were present from the beginning of each sorting task in Expt. II. In Expt. I, the subject had to remember the position that he had chosen for each of the piles; no order was given and there was no inherent 'natural' order like the numerical order which subjects spontaneously adopted in the

pilot experiment for Expt. II without marker cards. Compared with Expt. II therefore, both the stimulus material and the absence of markers combined to impose a greater load on memory in Expt. I and the interaction of dose and stimulus information could have arisen from an impairment of short-term retention and retrieval by the drug. This conclusion is supported by an earlier finding: nitrous oxide produced a systematic, dose-related impairment of short-term retention and retrieval in a continuous task which required oral responses to be made when sequences of numerals were presented visually (cf. Summerfield, 1964*a*). The difference in the interaction between stimulus information and dose of drug in the two experiments can therefore be regarded as indicating that task complexity was not equally well graduated by the measure of stimulus information in the two cases; it confirms arguments of other kinds that additional processes intervene in classification tasks like sorting playing cards which involve differential selection and suppression of aspects of the stimuli (cf. Posner, 1965.)

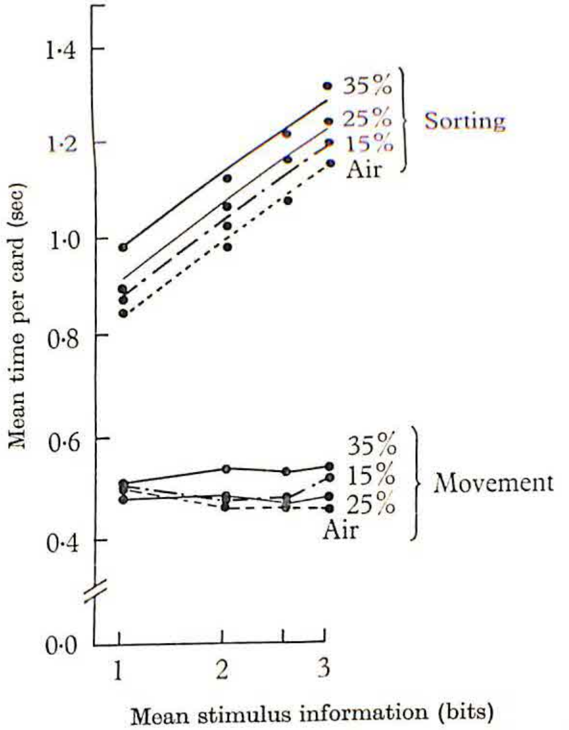


Fig. 2. Mean sorting time/card for four tasks (1, 2, 2·59, 3 bits/stimulus) at four dose levels of nitrous oxide in oxygen (0 %—air, 15, 25, 35 %), with times for movement only for comparison, in Expt. II.

Generally, there are three points in favour of communication models in studying effects of drugs on behaviour. First, sequential tasks are more readily discussed in terms of such models than by S–R theories, and it is these tasks which are most representative of every day skilled performances. Side effects of therapeutic drugs on industrial and other skills such as driving are an important practical problem and experiments on effects of drugs on sequential tasks are relevant to it. Secondly, information theory deals with the question of ensemble size, i.e. the effect of the

number of signals which might have occurred but did not. This factor is not discussed by S–R theories and they cannot at all readily be extended to include it. Thirdly, a communication model provides a useful framework for discussion of effects of drugs in terms of functional locus of action. As Broadbent (1959) has said, S–R theories make no attempt to relate intervening variables to physiological observations. Yet, for a proper understanding of the behavioural effects of drugs, it is essential to have a psychological framework into which pharmacological and physiological data can be fitted. None of the disciplines is at present sufficiently advanced for findings in one to be directly linked with those in any other. The communication model allows the nervous system to be described in terms of a block diagram like that of an electrical system. It gives a broad functional scheme into which details of structure can be fitted as new results are obtained. A complementary possibility is that of analysing processes involved in skilled performance by using drugs as experimental tools (Summerfield, 1961). If a drug selectively impairs one stage in a chain of functional mechanisms, the contribution of this stage to the performance of different tasks can be examined. Our findings can be represented in terms of a simple block diagram, similar to those discussed by Broadbent (1958), Welford (1960) and Treisman (1964), with a pre-central input stage, a central coding or 'choice' stage, and an effector stage arranged in sequence to form a channel, with storage systems for short-term and long-term memory. In these terms, our expectation that the drug would interact with the choice stage to decrease the rate of coding, in addition to any general reduction in transmission, was not confirmed. The drug, however, enabled movement components, on which it had no effect at the doses that were used, to be distinguished from other components involved in sorting. In Expt. II with numeral cards, each dose produced a constant reduction in speed at each level of choice, implying a constant lag at some stage earlier than choice or movement; this result, together with the result of Expt. I and of other experiments reviewed by Summerfield (1964*a*) implicate the input stages, in interaction with memory.

REFERENCES

ALLUISI, E. A. & MARTIN, H. B. (1958). An informational analysis of verbal and motor responses to symbolic and conventional arabic numerals. *J. appl. Psychol.* **42**, 79–84.

ALLUISI, E. A. & MULLER, P. F. (1958). Verbal and motor responses to seven symbolic visual codes: a study in S–R compatibility. *J. exp. Psychol.* **55**, 257–54.

BERRY, C., GELDER, M. G. & SUMMERFIELD, A. (1963). Information theory concepts in the experimental analysis of drug effects. *Bull. Brit. psychol. Soc.* **16**, no. 50, 3A.

BROADBENT, D. E. (1958). *Perception and Communication*. Oxford: Pergamon.

BROADBENT, D. E. (1959). Information theory and older approaches in psychology. *Proc. XVth int. Congr. Psychol., Brussels*, 1957, pp. 111–15. Amsterdam: North-Holland.

COCHRAN, W. G. & COX, G. M. (1957). *Experimental Designs*, 2nd ed. New York: Wiley.

COX, D. R. (1961). *The Role of Statistical Methods in Science and Technology*. London: Birkbeck College.

CROSSMAN, E. R. F. W. (1953). Entropy and choice time: the effect of frequency unbalance on choice-response. *Quart. J. exp. Psychol.* **5**, 41–51.

CROSSMAN, E. R. F. W. (1964). Information processes in human skill. *Brit. med. Bull.* **20**, no. 1, 32–7.

DEININGER, R. L. & FITTS, P. M. (1955) Stimulus-response compatibility, information theory and perceptual-motor performance. In *Information Theory and Psychology*, ed. H. Quastler. Glencoe, Ill.: Free Press.

LEGGE, D. (1965). Analysis of visual proprioceptive components of motor skill by means of a drug. *Brit. J. Psychol.* **56**, 243–54.
POSNER, M. I. (1965). Memory and thought in human intellectual performance. *Brit. J. Psychol.* **56**, 197–215.
STEINBERG, H. (1954). Selective effects of an anaesthetic drug on cognitive behaviour. *Quart, J. exp. Psychol.* **6**, 170–80.
STEINBERG, H. (1959). Effects of drugs on performance and incentives. In *Quantitative Methods in Human Pharmacology and Therapeutics*, ed. D. R. Laurence. Oxford: Pergamon.
SUMMERFIELD, A. (1961). Drugs as research tools in psychology: serial order and coding. *Acta psychol., Hague*, **19**, 778–80.
SUMMERFIELD, A. (1964*a*). Drugs and human behaviour. *Brit. med. Bull.* **20**, no. 1, 70–4.
SUMMERFIELD, A. (1964*b*). Codes and plans and organized behaviour. *Bull. Brit. psychol. Soc.* **17**, no. 57, 1–17.
SUMMERFIELD, A. & LEGGE, D. (1960). Perception and information theory. *Bull. Brit. psychol. Soc.* no. 40, 23–6.
SUMMERFIELD, A. & LEGGE, D. (1964). Information theory and perception. In *Readings in Psychology*, ed. J. Cohen. pp. 236–48. London: Allen and Unwin.
TREISMAN, A. M. (1964). Selective attention in man. *Brit. med. Bull.* **20**, no. 1, 12–16.
WELFORD, A. T. (1960). The measurement of sensori-motor performances: survey and re-appraisal of twelve years progress. *Ergonomics*, **3**, 189–230.

(*Manuscript received* 15 *April* 1964)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 267–273
Printed in Great Britain

PERSONALITY AND THE INVERTED-U RELATION

By D. W. J. CORCORAN
Medical Research Council, Applied Psychology Research Unit, Cambridge

Difficulties with the postulated inverted-U relationship between performance and arousal are discussed, with emphasis upon individual differences in level of arousal. Predictions concerning the behaviour of highly aroused and less aroused subjects are made and tested in two experiments by relating changes in performance associated with increased and decreased levels of arousal to introversion score. Introverts behaved as highly aroused subjects were expected to and extraverts as less aroused subjects.

Many workers have found the assumption of an inverted-U relation between performance and level of arousal to be useful. Some (e.g. Hebb, 1955; Malmo, 1959; Duffy, 1949, 1957) have used the relation as an aid to theory, while others (e.g. Freeman, 1940; Courts, 1942; Schlosberg, 1954; Stennett, 1957) have found that the relation fits their data. The assumption is, however, loose and ill-defined since with a U function direct prediction of the value on one axis from knowledge of the other is not always possible. It is the purpose of the present paper to show how this difficulty may be overcome and to demonstrate that some of the predictions the assumption yields may be valid.

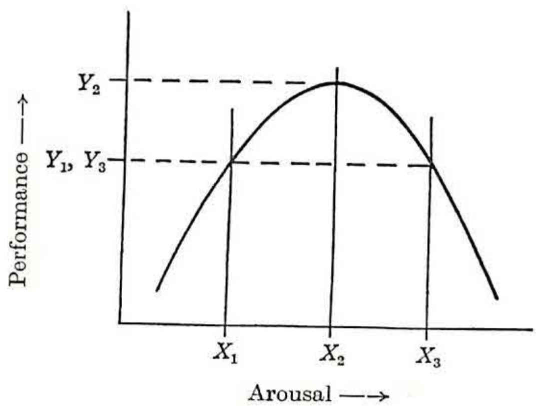


Fig. 1. Performance level Y_1, Y_3 would result from arousal levels X_1 or X_3. Given values Y_1, Y_3 it is possible to determine whether arousal level is at X_1 or X_3 by manipulating level of arousal and noting the directional change in performance.

An inverted-U relation implies that for any given value of performance except the optimal there will be two possible values of arousal. So that although level of performance is predictable, given level of arousal, level of arousal cannot be ascertained merely from knowledge of performance. However, there are instances in which the latter prediction is possible. Suppose level of performance was initially at Y_1 in Fig. 1, that with increased arousal performance rose to Y_2 and with a further increase dropped to Y_3 (where $Y_3 = Y_1$). Given only the value Y_1/Y_3, level of arousal could be either at X_1 or X_3. It is possible to decide which of the levels of arousal gave rise to performance level Y_1 either by increasing level of arousal or by decreasing it. If

arousal is lowered then Y_3 will increase, but Y_1 will decrease; similarly Y_1 will increase and Y_3 decrease if level of arousal is raised. Thus level of arousal can be predicted from level of performance, provided that the former is *varied* and the resulting performance level compared with that before level of arousal was changed.

Individual differences in arousal

The method suggested to determine level of arousal when level of performance only is given can be applied to discover which members of a group of subjects are operating at high levels of arousal and which at low levels. A group of subjects who are high in arousal may be considered to occupy a position along the abscissa of Fig. 2 somewhere to the right of the low arousal group. Although the following

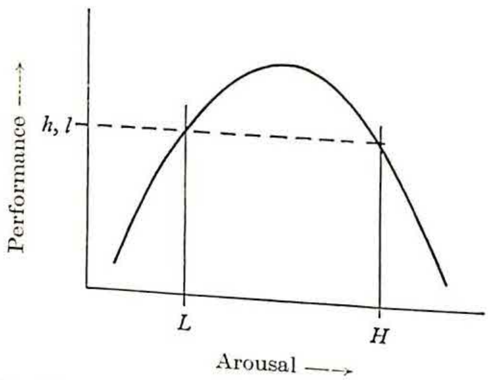


Fig. 2. Groups L and H both perform at h, l. By manipulating level of arousal it is possible to determine which group is at H and which at L.

arguments apply to any pair of positions along the abscissa we shall, for convenience, locate the high arousal group at H and the low at L, such that the level of performance of both groups is h, l. It is possible to discover which group has the high level of arousal either by decreasing the level of arousal of all subjects or by increasing it, since if level of arousal is decreased H's performance will improve by ascending the curve but L's performance will decline. Similarly, H's performance will decline and L's performance improve if the level of arousal of both is increased. The two groups may be differentiated by these methods no matter what positions they occupy along the arousal dimension, provided that H is to the right of L. Thus, for example, if H were at the optimal position and L somewhat lower, L would improve more than H if arousal were increased and decline more if arousal were decreased.

The following experiments were designed to test the validity of the foregoing argument. These experiments required (1) a method of changing arousal level, and (2) a technique for classifying subjects according to whether they would be likely to be operating at a high or low level of arousal. Deprivation of sleep was used to manipulate level of arousal (Corcoran, 1964*a*). The second requirement was met by relating the results to scores of introversion, since there is some evidence indicating that level of arousal may be related to degree of introversion (Shagass & Kerenyi, 1958; Claridge, 1961; Corcoran, 1962, 1964*b*; Colquhoun & Corcoran, 1964).

Experiment I

Method

Task, subjects and procedure

The 'five-choice serial-reaction task' (Leonard, 1959) was used. The subject was seated in a sound proof cubicle, facing a display of five bulbs arranged in a pentagon on a board slightly inclined from the vertical. On the table before him was a board on which were inset five circular brass disks, each $1\frac{1}{2}$ in. in diameter. The disks were set in spatially equivalent positions to the five bulbs. The subject was instructed to tap the disk corresponding in position to the bulb which was alight at the time. By tapping the disk, the light was switched off and another bulb lit. The sequence in which the bulbs light repeats itself after 100 taps, and was thus effectively unpredictable. The apparatus was designed to yield three separate measures of performance: hits, the number of times the subject tapped the disk corresponded to the light; errors, the number of times an incorrect disk had been tapped; and gaps, the number of times an interval of $1\frac{1}{2}$ sec. had elapsed between responses.

An 'incentive' condition (HM) was used as a method of increasing level of arousal. Sleep deprivation was used as a de-arouser. The incentive was created by the following instructions: 'There is no time limit to this task; when you can finish depends how well you do. You have to make 3000 correct taps, and as soon as you have done this you can stop. You will be told when this figure has been attained. Errors do not count towards the total'. In the LM condition the subjects were simply requested to carry out the task for half an hour; they would usually make about 2000 hits during this time. All subjects worked for half an hour irrespective of number of hits.

Twelve naval ratings were tested under each of the following four conditions, according to a 2×2 latin-square design: (*i*) after normal sleep on the previous night, and under normal (low incentive) conditions (S/LM); (*ii*) after normal sleep, and after being given the instructions designed to result in a high level of incentive (S/HM); (*iii*) after losing sleep on the previous night and with low incentive (NS/LM); (*iv*) after loss of one night's sleep and with high incentive (NS/HM).

The Heron scale of introversion-extraversion (Heron, 1956) was administered to all subjects by laboratory staff unconnected with the present work. The scores on the Heron test showed that the subjects were on average slightly extraverted, having a mean score of 2·6 with an S.D. of 1·68.

Results and discussion

Spearman rho correlations were computed between introversion-extraversion and twenty-four measures of performance. The measures and the correlations are recorded in Table 1. In Table 2 are recorded the data relevant to the effects of increased incentive.

Table 1. *Correlations between introversion and measures of performance at the five-choice serial-reaction task, after loss of sleep (NS) and normal sleep (S), under conditions of low (LM) and high (HM) motivation*

Comparison no.	Condition	Gaps	Hits	Errors
1	S/LM	−0·29	0·68*	+0·11
2	S/HM	−0·16	0·45	−0·23
3	NS/LM	−0·72*	0·79*	−0·01
4	NS/HM	−0·51	0·18	+0·16
5	S/LM-S/HM	−0·26	0·43	+0·03
6	NS/LM-NS/HM	−0·63*	0·59*	+0·12
7	S/LM-NS/LM	+0·76*	0·38	−0·05
8	S/HM-NS/HM	+0·51	0·02	+0·05

* $P < 0.05$

The results of statistical significance in Table 1 may be summarized as follows. (*a*) Introverts made more hits than extraverts under S/LM and NS/LM conditions and scored fewer gaps under NS/LM conditions. (*b*) Extraverts deteriorated more from loss of sleep in terms of gaps under LM conditions, but improved more from the incentive, both in terms of gaps and hits. The two latter findings were significant only under the NS condition (see Table 2).

Table 2. *Performance at the five-choice serial-reaction task. Mean scores of the six most introverted subjects (I) and six most extraverted (E) after loss of sleep (NS) and normal sleep (S) under conditions of low (LM) and high (HM) motivation*

	Gaps					Hits				
	I			E			I			E	
	LM	HM		LM	HM		LM	HM		LM	HM
S	32·8	24·5	S	92·8	65·8	S	2932·8	2981·8	S	2205·8	2640·8
NS	107·0	33·8	NS	562·8	300·3	NS	2197·6	2393·8	NS	1124·2	2186·5

The results can be condensed into three main findings. (i) Introverts performed the task rather better than extraverts. If we accept the terminology and assumptions presented in the introduction, we may deduce from this finding that the task situation as a whole proved to be a relatively low in arousal value, since only under such conditions would a group operating at a high level of arousal be expected to perform better than one operating at a low level. (ii) The arousing procedure probably benefited extraverts more than introverts. This finding proved to be reliable only under NS conditions in the present experiment, but since other experiments have shown the effect after normal sleep (Corcoran, 1962), the generalization is probably warranted. (iii) Extraverts were probably the most affected by the de-arousing procedure. This result was found to be statistically significant only under LM conditions in terms of gaps.

The effect upon performance of changing level of arousal was shown to be about what one would expect on the basis of the inverted-U relationship. The less aroused subjects (if we can equate these with the extraverts) were more affected by lowering arousal, but improved relative to the highly aroused subject when level of arousal was increased. The results were far from conclusive, however, and the following further experiment was therefore carried out.

Experiment II

The task used in Expt. I was better performed by introverts. It was deduced from this finding that the nature of the task and the testing situation as a whole resulted in a suboptimal level of arousal, since under no other condition would a group high in arousal be expected to perform better than one which is low. Similarly, if a group is operating at a super-optimal level of arousal those subjects within the group who are low in arousal should perform the best, since they will be closer to the optimum. The aim of the following experiment was to investigate the effect of a de-arousing procedure upon groups who, it was hoped, were at post-optimal levels. On the basis of the Yerkes–Dodson principle, it would seem reasonable to choose a difficult task since these may be characterized by a low optimal level of arousal. A difficult task

was therefore chosen, and the subjects were also put under stress, so that the low optimum level of arousal plus the arousing situation might be expected to result in a group who were operating at a high level of arousal. The effect to be expected from a reduction in arousal in such a situation is clear: the more aroused subjects should show more improvement than the less aroused.

Method

Task, subjects and procedure

A modified triple tester was used for this experiment. This apparatus consisted of a revolving brass drum with a celluloid cover. The cover was perforated with circular holes of $\frac{1}{4}$ in. diameter, arranged in such a way as to leave a passage free from holes. This passage wound spirally from left to right following a winding course. A pointer, controlled by means of a steering wheel, had to be kept within the passage as the drum revolved. When the pointer left the passage, it immediately made contact through the holes with the surface of the brass drum. This completed a circuit which included a counter and an error was scored. The error score thus indicated the approximate proportion of the time which the pointer was off track.

Subjects were tested under three separate conditions. (1) The drum was set to revolve at such a speed that the trial was completed in 40 sec.; the view of the oncoming track was restricted; each error was punished by a blast of about 100 db. white noise through earphones. (2) Conditions were identical to those described in 1, but three full sessions were allowed as practice before the first scored session. (3) Conditions were identical to those described in 1, but only the click of the error counter was heard through the earphones when an error occurred. These three conditions made up part of a larger investigation, which included other experimental conditions irrelevant to the present discussion.

A session consisted of ten trials; in each trial the pointer had to be kept in a position from the start of the passage on the left, through three and a half revolutions of the drum to end on the right. Three sessions were conducted on three consecutive mornings between 7·30 a.m. and noon.

Eighteen subjects were tested during a 60 hr deprivation of sleep: on the first morning (Day 1) they worked after sleeping normally on the previous night; on the second morning (Day 2) they had lost their normal sleep on the previous night; on Day 3 they worked after two nights' loss of sleep. Six subjects carried out the task under each of the three conditions (1), (2) and (3).

Results and discussion

The error scores were converted into normal deviates in order to make the scores more comparable under the three sets of conditions. Table 3 presents the introversion and extraversion scores of the subjects employed in the present experiment.

Table 3. *Introversion and extroversion scores: mean scores and ranges for subjects in Expt. II* (n = 18)

Group		n	Mean	Range
NS	*I*	9	6·2	4–9
	E	9	2·0	1–3

The converted tracking scores are presented in Table 4 in which the changes in performance over the 3 days can be seen. Extraverts show a consistent decline over the sleep deprivation, whilst introverts improve to an almost equal extent. These trends in performance differ significantly when assessed by the Mann–Whitney test ($U = 8$, $P = 0{\cdot}002$).

These were not exactly the results expected, since the task was designed to be hyper-arousing to all subjects, so that all were expected to show an improvement in performance but the more aroused to the greater extent. The less aroused (if we can

equate these subjects with the extraverts) deteriorated rather than improved with loss of sleep. The results therefore suggest that the less aroused subjects were initially at a suboptimal level of arousal.

Table 4. *The converted scores of extraverts ($n = 9$) and introverts ($n = 9$) on each day of sleep deprivation for the tracking task*

	Day 1		Day 2		Day 3	
	Mean	Range	Mean	Range	Mean	Range
Introverts	−0·21	−1·96 to +1·59	+0·40	−0·86 to +1·88	+0·70	−0·42 to +1·43
Extraverts	+0·19	−1·14 to +1·22	−0·41	−1·44 to +1·10	−0·72	−1·71 to +0·46

The results may in fact indicate that a phenomenon rather like that shown in Fig. 2 was operating. In discussing the implication of the situation illustrated in Fig. 2 it was pointed out that the knowledge of level of performance alone was insufficient to determine level of arousal and that arousal level had to be changed in order to determine at which of the points (H or L) the subject had been operating. In the present experiment, initial level of performance was not very different and by manipulating level of arousal it was possible to ascertain which 'type' operated at H and which at L.

Conclusions

It was suggested in the introduction that information could be gained about the level of arousal at which a subject was operating by increasing or decreasing his level of arousal and noting the direction of the corresponding change in performance. The experiments were an attempt to demonstrate that this could be done. However, the results demonstrated the validity of the argument only to the extent that the groups used as external criteria were valid ones. In other words the argument has been shown to be valid provided that introverts are in general more highly aroused than extraverts. Although the results of many experiments showing differences between introverts and extraverts can be easily *interpreted* to be due to different levels of arousal, this hypothesis is far from proven.

Assuming, however, that the criterion groups were valid, the experiments showed (*a*) that the performance of the less aroused subject deteriorates when the general level of arousal is decreased, (*b*) that the performance of the more aroused subject is less affected than that of the less aroused subject when arousal generally is decreased and may even improve if the initial level of arousal is past the optimum.

References

Claridge, G. S. (1961). Arousal and inhibition as determiners of the performance of neurotics. *Brit. J. Psychol.* **52**, 53–4.

Colquhoun, W. P. & Corcoran, D. W. J. (1964). The influence of time of day and social isolation on the relationship of temperament to performance. *Brit. J. soc. clin. Psychol.* **3**, 226–31.

Corcoran, D. W. J. (1962). Individual differences in performance after loss of sleep. Unpublished Ph.D. Thesis, University of Cambridge.

Corocoran, D. W. J. (1964*a*). Changes in heart rate and performance as a result of loss of sleep. *Brit. J. Psychol.* **55**, 304–14.

Corcoran, D. W. J. (1964*b*). The relationship between introversion and salivation. *Amer. J. Psychol.* **77**, 298–300.

Courts, F. A. (1942). Relationships between muscular tension and performance. *Psychol. Bull.* **39**, 347–67.

Duffy, E. (1949). A systematic framework for the description of personality. *J. abnorm. soc. Psychol.* **44**, 175–90.

Duffy, E. (1957). The psychological significance of the concept of arousal or activation. *Psychol. Rev.* **64**, 265–75.

Freeman, G. L. (1940). The relation between performance level and bodily activity level. *J. exp. Psychol.* **26**, 602–8.

Hebb, D. O. (1955). Drives and the CNS (Conceptual Nervous System). *Psychol. Rev.* **62**, 243–54.

Heron, A. (1956). A two part personality measure for use as a research criterion. *Brit. J. Psychol.* **47**, 253–41.

Leonard, J. A. (1959). Five choice serial reaction apparatus. A.P.U. 326. Med. Res. Council Appl. Psychol. Res. Unit, Cambridge.

Malmo, R. B. (1959). Activation: A neuropsychological dimension. *Psychol. Rev.* **66**, 367–86.

Schlosberg, H. S. (1954. Three dimensions of emotion. *Psychol. Rev.* **61**, 81–5.

Shagass, C. & Kerenyi, A. B. (1958). Neurophysiological studies of personality. *J. nerv. ment. Dis.* **126**, 141–7.

Stennett, R. G. (1957). Performance level and level of arousal. *J. exp. Psychol.* **54**, 54–61.

(*Manuscript received* 11 *April* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 275–283
Printed in Great Britain

DISCRIMINATION OF AVOIDABLE AND UNAVOIDABLE SHOCK

By ROBERT BROWN
*Department of Psychology, University of Exeter**

Rats were given standard avoidance training in a two-compartment shuttle-box. Thereafter shock was only signalled if the subject was in one of the distinctively coloured compartments at the start of a given trial, while non-signalled shock was presented in the other compartment. Subjects showed a significant preference for this compartment on several measures. Half of the subjects received a relatively higher shock intensity during this latter part of the experiment, and this had the effect of improving avoidance performance and increasing the subjects' preference for the signalled shock situation. These findings are discussed in terms of the relative conditioned aversiveness of various stimuli present in the situation.

When an originally neutral stimulus is repeatedly paired with one which produces a reliable response, it acquires the capacity to control behaviour. Traditionally, a warning signal which precedes an electric shock is held to take on some of the properties of the shock itself. It is supposed to become in some way 'fear-arousing' or, in more empirical terms, 'secondarily aversive' (Schoenfeld, 1950; Miller, 1951; Dinsmoor, 1954), since its removal or termination can act as a reinforcing event. However, this traditional view has been questioned recently by several investigators. Keehn (1959) has shown that animals will learn to postpone an electric shock, but not a signal which precedes the shock. His inference was that animals should learn to postpone the signal in like manner if the signal is in some way 'similar' to the shock. The finding that they do not led him to suggest that the signal functions entirely as a discriminative stimulus, and that any reinforcing capacity it may possess is incidental.

A more direct attack on the problem has been made by Lockard (1963). Rats were allowed to choose between a situation in which an unavoidable shock was presented alone, and one in which the unavoidable shock was preceded by a warning signal. Significant preferences were shown for the signal situation. The implication is that the signal acquires certain properties which cannot be classed as aversive ones. Although it may not act as a positive reinforcer for any clearly identifiable operants, it is more of a 'satisfier' than an 'annoyer' in Thorndike's sense. One might expect that the animals would have shown an even more marked preference for the signal situation if they had been able to *avoid* the shock on receipt of the signal. On the other hand, avoidance of shock is usually accomplished by escape from signal; an animal would have to learn to terminate a danger signal when the signal was present, but to 'seek out' the signal when it was absent. Such a mode of behaviour might appear paradoxical. How could termination of a signal act as a reinforcer if its absence were nevertheless 'unpleasant'?

The paradox is resolved in so far as 'signal-seeking' behaviour becomes at least biologically advantageous if absence of signal sets the occasion for the delivery of

* The work reported here was carried out while the author was at the Department of Psychology, University of Bristol.

unavoidable shocks. Empirical evidence on this point is somewhat sparse, however, and the present experiment was carried out in an attempt to clarify the issue. The choice technique was basically similar to that of Lockard (1963) and apparently first used by Akhtar (cf. Mowrer, 1960, pp. 194–5). Akhtar gave four rats one trial per day in a Miller–Mowrer shuttle-box. A warning signal preceded shock in only one of the two compartments, and the prescribed avoidance response was a leap into the air. Shock was presented alone in the other compartment. All four rats showed a preference for the signal compartment, but only two learned to avoid shock.

There are four main differences between this procedure and the one adopted in the present study.

(i) Each animal had three sessions of fifty trials rather than only one choice trial per day; this permitted a more detailed analysis of a given subject's behaviour over a longer period of time.

(ii) Prior to the choice procedure, two fifty-trial sessions of standard avoidance training were given. This meant that subjects had been escaping from the warning signal rather consistently without any control over its onset. On the other hand, in later choice sessions an animal could choose whether to receive a signal or not, and thus had at least indirect control over signal onset as well as offset. This initial training also had the practical advantage of shaping the specified escape and avoidance responses, but perhaps most importantly, it provided an essentially symmetrical baseline situation against which performance in the asymmetrical choice situation could be compared.

(iii) Avoidance responses, escape responses and 'choice-of-compartment-responses' were all defined by the same criterion: any behaviour which resulted in a barrier-crossing. During a choice session, then, consistent avoidance could only be maintained if the animal returned equally consistently to the signal compartment. If avoidance and escape responses had been defined as in Akhtar's study (a leap into the air), then a subject might well have 'chosen' to receive the signal for many successive trials; but since this would simply have meant *staying* in the signal compartment, the choice or choices would have been covert rather than overt. By means of our procedure, the subjects were forced to move between compartments and thus experience the stimulus events connected with each; and if they 'chose' to receive two or more successive trials in a given compartment, such choices were observable as return barrier-crossing responses.

(iv) The final difference concerns the intensity of shock used. Initial avoidance training was the same for all subjects; 0·3 mA. shock was used. For the choice sessions, half the subjects continued with 0·3 mA., but half were shifted to 0·9 mA. Level of motivation should be a relevant variable if, as Mowrer (1960) has argued, responses which produce warning signals are in some way analogous to the observing responses studied by Wyckoff (1952) and Prokasy (1956). For example, Wehling & Prokasy (1962) have shown that increased food deprivation leads to a higher level of observing behaviour in a food reinforcement situation. The parallel here is that a higher level of shock might increase the probability of occurrence of signal-producing responses although it might also increase any secondarily aversive properties of the signal.

Method

Subjects

Sixteen male hooded rats, aged between 90 and 115 days at the start of the experiment, served as subjects. They were housed in cages of four and had access to food and water at all times.

Apparatus

The experimental chamber was a Miller–Mowrer type shuttle-box, length 36 in., width 6 in., height 9 in., with two compartments. The end walls were plywood painted a flat grey, the side walls and ceiling were of Perspex sheet, and the floor consisted of transverse $\frac{1}{8}$ in. diameter steel rods spaced $\frac{1}{2}$ in. centre to centre. An aluminium partition separated the box into equal halves; access between them was through a 4 in. square hole above a barrier $1\frac{1}{2}$ in. high. The upper part of this barrier was a copper tube mounted on bearings which successfully prevented the animals from 'perching'. The two compartments were made distinctive by placing painted card panels flush with the outside of the Perspex side walls, and painted faeces trays 1 in. below the grids. (For one half, the panels were black and the tray black; for the other half the tray was white and the panels coloured black and white in alternate 2 in. vertical stripes.) Illumination was provided by a 5 W. bulb placed centrally over each compartment, and the animal was observed through a 45° mirror over the bulbs.

The auditory warning signal was produced by activating a buzzer which was spring-mounted centrally outside the chamber. It raised the noise level in the chamber from 28 ± 2 db. to 65 ± 2 db.

Shock was applied to the grids and barrier from a Foringer constant current a.c. stimulator. The pattern of shock on the grid bars was changed approximately 7 times per sec. by an electro-mechanical scrambling device.

A system of relays and timers programmed the presentation of stimuli. Stimulus and response events were recorded on an Evershed pen recorder. Programming and recording equipment was sound deadened as far as possible and housed in an adjacent room. The only non-automatic feature of the equipment was a silent changeover switch operated manually by the experimenter when the subject crossed the barrier.

Procedure

Each animal spent a 1 hr. session in the shuttle-box on each of seven successive days at the same time every day (± 20 min.), between the hours of 10 a.m. and 4 p.m. The subjects were transported between the home cage and the experimental chamber in a small wooden box, which also had a removable floor and wall to ensure uniform insertion and removal. Each session consisted of 50 min. experimental time preceded and followed by 5 min. adaptation periods. Subjects were allotted to experimental treatments on the basis of free-feeding weight and age, prior to any experimentation.

The first session was solely to adapt the animal to the apparatus; no signals or shocks were presented. During the second session, the buzzer was turned on at the beginning of each minute and terminated either after 20 sec. or after a barrier-crossing, whichever was sooner. This provided a basis for discarding subjects if they showed a tendency to learn to terminate the buzzer in its own right.

Avoidance training was carried out in the third and fourth 'avoidance' sessions (A 1 and A 2). Subjects received 50 trials per session, with a regular 1 min. intertrial interval. A delayed conditioning procedure was used, with a 5 sec. CS–US interval. If a barrier-crossing response was made during this 5 sec. period, the buzzer was terminated immediately and shock omitted for that trial; such a response thus constituted an avoidance response. A response occurring after shock onset terminated shock and buzzer simultaneously, and was thus an escape response. Subjects were free to make intertrial responses (ITR) at any time during the sessions. Shock intensity was set at 0·3 mA. (short circuit current) for all subjects.

For the remaining three 'choice' sessions (C 1, C 2 and C 3) shock intensity was raised to 0·9 mA. for half the subjects, but stayed at 0·3 mA. for the other half, and the choice procedure was introduced for all subjects. This procedure differed from the two preceding avoidance sessions in the following way. One of the compartments was defined as the 'signal' compartment

and the other as the 'non-signal' compartment. If the subject happened to be in the signal compartment at the start of a trial, it received exactly the same sequence of stimuli as before; in fact, a conventional avoidance trial. If the subject was in the non-signal compartment at the start of the trial, however, shock was presented simultaneously with the buzzer, and clearly only escape responses were possible. (For half of the subjects in each shock intensity subgroup the black compartment was the signal compartment and the striped one the non-signal compartment; these conditions were reversed for the other subjects.) Thus in order to avoid consistently, the subject had to be present in the signal compartment at the start of a trial and make an ITR which returned it to that compartment ready for the next trial. ITR were recorded and classified in terms of whether they removed the subject from the signal to the non-signal compartment, or vice versa. The time at which they occurred during the 1 min. cycle of each trial was also noted.

Results

None of the subjects had to be discarded for learning to terminate the buzzer in session 2. Two groups of eight subjects were thus formed, Group H (high-shock) and Group L (low-shock). The groups are hereafter referred to as H and L, although in fact they only received different shock intensities from session C1 onwards. Similarly, compartments are called 'signal' and 'non-signal' throughout, although these labels only have empirical significance when referring to 'choice' sessions.

In sessions A1–2, when all subjects were treated alike, there were no significant differences on any performance scores between the two groups. Thus it was concluded that allocation to subsequent treatments had in fact been random. As expected, avoidance performance improved over these two sessions, and frequency of intertrial responding decreased. No preference was shown for either compartment in terms of the preference measures discussed below.

With the introduction of the choice procedure in sessions C1–3, several major changes appeared. For Group L, mean avoidance performance on C1–3 (43·6%) was significantly poorer than on A1–2 (73·6%) (sign test, 2-tail $P = 0{\cdot}008$). Group H showed a smaller non-significant performance decrement (69·6–63·0%; sign test, 2-tail $P = 0{\cdot}73$). Thus Group H made significantly more avoidances than Group L over sessions C1–3 (Mann–Whitney $U = 9$, 2-tail $P = 0{\cdot}014$).

However, these performance changes are not wholly meaningful until it is remembered that avoidance was possible only if the subject was occupying the signal compartment at the start of trials in sessions C1–3. Fig. 1 shows the mean percentage of trials per session which were begun with subjects occupying the signal compartment. Both groups showed a significant tendency to be present in the signal compartment at the start of trials on sessions C1–3, as compared with their essentially chance preference on sessions A1–2, when all trials were signalled. Each subject showed an increased preference. Moreover, Group H subjects were present in the signal compartment at the start of 79·4% of the C1–3 trials, as against 63·8% for Group L. These group means were significantly different ($U = 6$, 2-tail $P = 0{\cdot}004$). It is clear that these two percentages set the maximum avoidances possible for the two groups. When the total avoidances are expressed as percentages of *these* maxima, it becomes apparent that Group H subjects avoided on 73·8% of the occasions on which avoidance was possible, while the corresponding figure for Group L is 66·8%. This difference between the groups suggests that the higher shock intensity had a facilitating effect on avoidance performance *per se*, although the difference does not

reach an acceptable level of significance ($U = 16$, 2-tail $P > 0{\cdot}1$). It can be seen from these revised performance scores that the decrements in absolute avoidance performance which were produced by the introduction of the choice procedure were mainly a result of the subjects simply not being in a position to avoid.

Fig. 1 also shows another measure of compartment preference, the mean percentage of intertrial time spent in the signal compartment. When the two preference measures are compared, the functions for Group H are reasonably similar in shape.

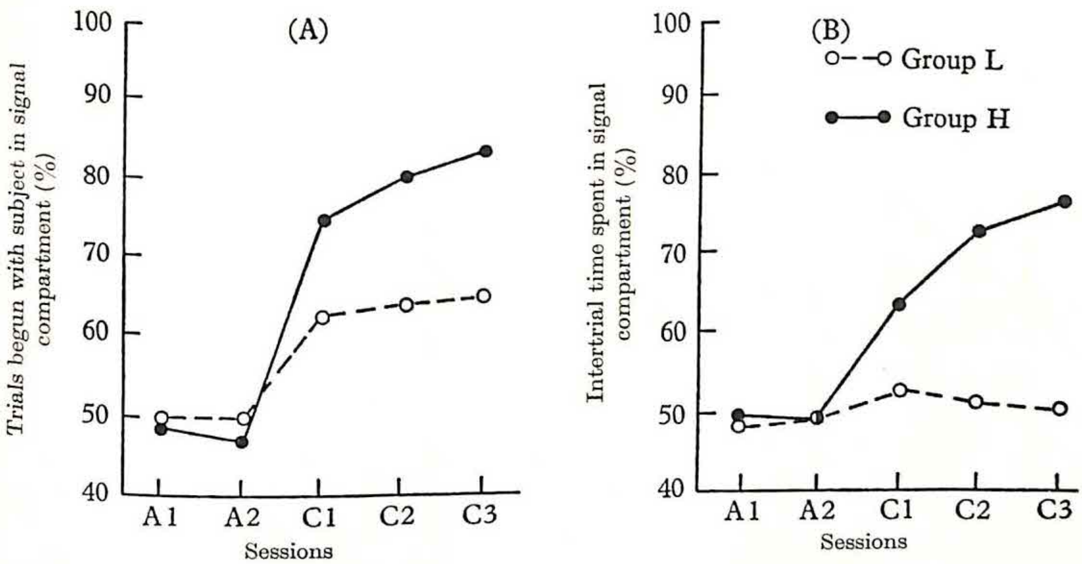


Fig. 1. Two measures of compartment preference. (A) shows the mean percentage of trials per session commenced with the subject occupying the signal compartment, for the two groups. (B) shows the mean percentage of intertrial time spent in the signal compartment. The 50 % point on each graph represents equal preference for the two compartments.

On the other hand, Group L showed little more than chance preference for remaining in the signal compartment, in spite of the fact that they were present there with greater than chance frequency at the start of the trials of sessions C1–3. In fact, Group H showed a net increase in time spent in the signal compartment of 21·6 % (mean of C sessions minus mean of A sessions; sign test, 2-tail $P = 0{\cdot}008$), while Group L showed a non-significant increase of only 2·5 %. This discrepancy between the two preference measures for Group L becomes understandable when the ITR data are examined.

Fig. 2 shows ITR as a function of their occurrence in the twelve successive 5 sec. periods within a 'trial'. Each histogram was constructed as follows. First, an ITR was classified as 'functional' if it was instrumental in returning a subject to the signal compartment, and 'non-functional' if it was in the opposite direction. All of the data from trials relevant to a given histogram were pooled, e.g. all of the functional ITR made by Group H subjects on all of the trials in sessions C1–3. Each total thus obtained was split up into twelve subtotals, according to which of the 5 sec. periods the responses occurred in. These twelve subtotals are represented by the heights of the twelve columns of each histogram. The extreme left-hand column of any histogram represents the first 5 sec. period of a trial, and the extreme right-hand column

represents the last 5 sec. period. (The 'first' 5 sec. period of each trial was the one which began with signal onset, or signal-plus-shock onset). Finally, the histograms were drawn on the same scale, i.e. each histogram represents data from a mean A or C session.

Ignoring the temporal distribution of responses, it can be seen that on sessions A1–2 each group made approximately equal numbers of functional and non-functional ITR: they showed no tendency to cross to one compartment rather than to

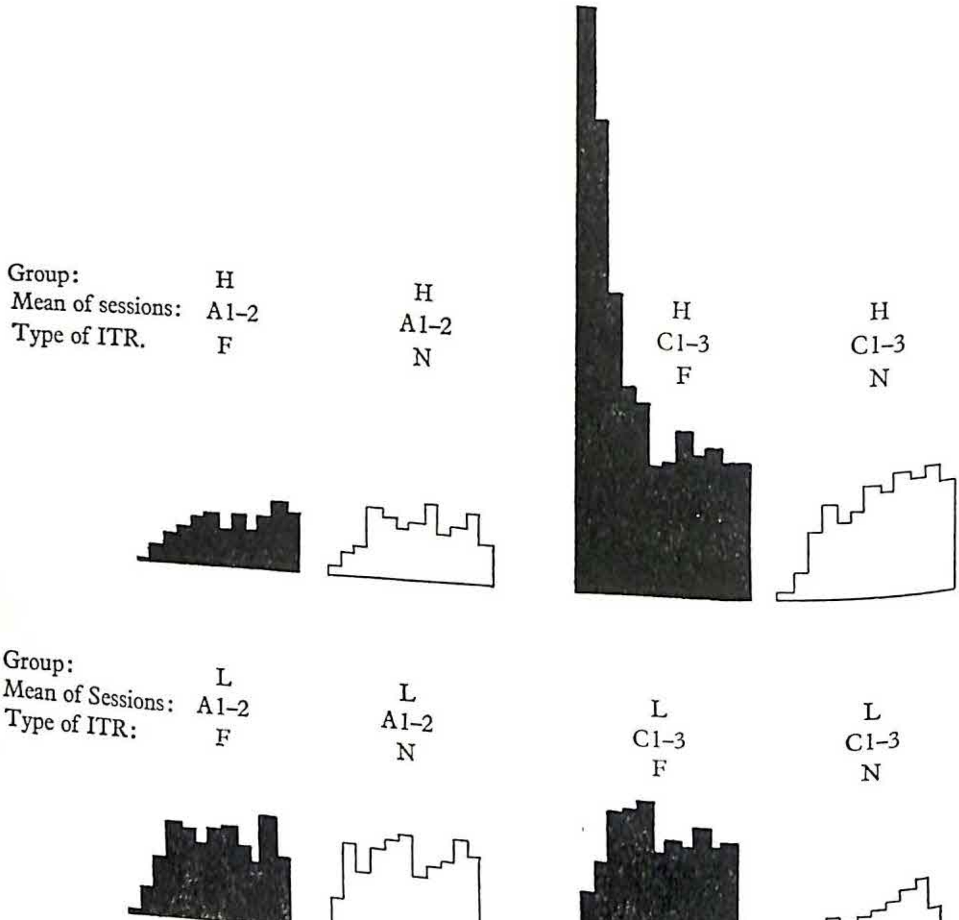


Fig. 2. Distribution of functional (F) and non-functional (N) intertrial responses (ITR) according to their temporal location in the 1 min. trial cycle. The height of each column represents response frequency, and position on the baseline represents the temporal location of the responses. The extreme left-hand column of each histogram corresponds to the first 5 sec. period of a trial, and the extreme right-hand column corresponds to the last 5 sec. period. (See text for further explanation.)

the other. When the compartments differed in 'significance' on sessions C1–3, all eight of the Group H subjects and seven of the Group L subjects made more functional than non-functional responses. Group H subjects made an average of 30·3 more functional than non-functional ITR on a mean C session, as against 14·2 for Group L subjects. These means were significantly different ($U = 6$, 2-tail $P = 0{\cdot}004$). Thus both groups made more responses which returned them to the signal compart-

ment than responses which took them out of it, on sessions C1–3, and the higher shock intensity made this pattern of behaviour more marked.

When the temporal location of these two types of ITR is examined, it is clear that all the frequency distributions are reasonably similar in shape except the one representing functional ITR for Group H on sessions C1–3. With this exception, the general trend is for ITR of either type to occur with relatively greater frequency towards the end of a trial. This means that for Group L, on sessions C1–3, the time spent in both compartments would be approximately equal, as was the case. However, for Group H, the bulk of the functional ITR occurred early in the trial, and most of the non-functional ITR occurred late in the trial; hence the net result of considerably more time spent in the signal compartment. (There was a slight tendency for Group L responses to approximate this patterm, which accounts for the slight, but non-significant increase in time spent in the signal compartment.)

Discussion

The clearest outcome of the experiment is that Akhtar's (Mowrer, 1960) preliminary observations have been confirmed; given a choice between a situation in which shock is signalled and avoidable and one in which it is not signalled and hence merely escapable, rats choose the former with some consistency. Moreover, the stronger the primary aversive stimulus, the greater the preference for the signal situation. It might appear then that the signal *is* non-aversive, and perhaps wholly discriminative as Keehn (1959) has suggested. Certainly, the subjects make a number of *discriminations* throughout the training. On sessions A1–2 a discrimination is formed between presence of signal and its absence—since avoidances increase and ITR decrease. On choice sessions a further discrimination is made between compartments; both groups are present in the signal compartment at the start of any given trial with greater than chance expectancy, and direction and point of occurrence of ITR contribute jointly to the significantly greater time spent in the signal compartment by Group H. Thus it is perfectly justifiable to claim that presence-of-signal and absence-of-signal act as a pair of discriminative stimuli, but to deny that these stimuli may also possess reinforcing properties would seem to be unwarranted. At least for positive reinforcement training, Dinsmoor (1950) has shown that cue and reinforcing functions develop concurrently.

While not denying that discriminations are present, then, the present findings can be interpreted in a way which attempts to identify the *bases* for these discriminations. The first assumption is simply that behaviour is normally homeostatic. In a situation involving aversive stimulation, behaviour which removes relatively aversive stimuli (and replaces them with relatively less aversive ones) will increase in probability of occurrence. Secondly, we assume that neutral stimuli become aversive by virtue of their temporal contiguity with the onset of primary aversive stimuli. These assumptions generally follow the positions adopted by Schoenfeld (1950) and Dinsmoor (1954) among others. There are thus two classes of stimulus-compound which can precede shock onset in the choice situation, and one which is never followed directly by shock. Because of the different probabilities of shock onset following these three compounds, they can be ordered in terms of relative conditioned aversiveness. The

compound which is *always* followed by shock onset consists (*a*) of proprioceptive stimuli arising from non-avoidance behaviour, along with (*b*) cues from being present in the signal compartment when (*c*) the exteroceptive warning signal is present. This compound (S1) would be relatively most aversive because of its inevitable association with shock. At the other end of the scale is compound S3 which is *never* directly paired with shock: this arises from the same behaviour in the same compartment in the *absence* of the signal Finally there is compound S2, which should be more aversive than S3 but less aversive than S1, and is the net result of non-avoidance behaviour in the non-signal compartment. This compound *is* sometimes followed by shock, but not in any immediate or consistent way.

A barrier-crossing in the presence of the signal thus changes the pattern of stimulation from S1 to S2. By reducing the relative aversiveness of stimulation, such a response increases in probability of occurrence. In other words, fairly stable avoidance of shock develops. By making a further barrier-crossing, our so-called 'functional' ITR, compound S3 is substituted for S2, and by the same token, these responses should likewise increase in frequency. Because there is considerable similarity between the three compounds, this 'ideal' pattern of behaviour is only approximated—but nevertheless this formulation provides an explanation for 'signal-seeking' behaviour without recourse to any but the simplest of behavioural principles.

If we make the third reasonable assumption that a more intense primary aversive stimulus leads to greater conditioned stimulus aversiveness, then the difference in performance between Groups H and L can also be accounted for. At a higher shock intensity, both S1 and S2 should increase in aversiveness, but S3 should increase negligibly, if at all, since it is never paired directly with shock. Therefore the differential aversiveness of non-avoidance behaviour in the two compartments is relatively greater for Group H, and hence the stronger preference shown for the signal compartment by this group between trials.

While the animals are clearly engaging in a form of 'observing behaviour' at a descriptive level, it appears that such behaviour can be analysed quite adequately in terms of well-established reinforcement principles. Certainly the results offer no support for Keehn's (1959) contention that warning signals are wholly discriminative in function. As Church (1963) has noted, aversive stimuli are usually defined in terms of the escape operation—and since, in conventional avoidance training, the subject is obviously escaping from the signal, then it is only natural that interpretations should have stressed the secondary aversive nature of the signal. On the other hand, in Sidman avoidance, as used by Keehn, the main explanatory burden carried by the aversive stimulus is its capacity to *punish* and thereby suppress non-avoidance behaviour. In both Keehn's experiment and the present one the signal is not aversive on the latter operational definition; but it is conceptually and empirically incorrect to let this finding outweigh the fact that the signal *is* aversive on the former definition. It may be unfortunate, but nevertheless true, that our theoretical concepts are sometimes shaped by the experimental situations we adopt.

In conclusion we would argue that our results support Dinsmoor's (1950) finding that the discriminative and reinforcing functions of a stimulus are inseparable. Even a food-pellet is discriminative in that it 'sets the occasion' for eating; but no one would deny that it is also reinforcing to a hungry rat.

References

Church, R. M. (1963). The varied effects of punishment on behavior. *Psychol. Rev.* **70**, 369–402.

Dinsmoor, J. A. (1950). A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *J. exp. Psychol.* **40**, 458–572.

Dinsmoor, J. A. (1954). Punishment: I. The avoidance hypothesis. *Psychol. Rev.* **61**, 34–46.

Keehn, J. D. (1959). The effect of a warning signal on unrestricted avoidance behaviour. *Brit. J. Psychol.* **50**, 125–35.

Lockard, J. S. (1963). Choice of a warning signal or no warning signal in an unavoidable shock situation. *J. comp. physiol. Psychol.* **56**, 526–30.

Miller, N. E. (1951). Learnable drives and rewards. In *Handbook of Experimental Psychology*, ed. Stevens, S. S. New York: Wiley.

Mowrer, O. H. (1960). *Learning Theory and the Symbolic Processes.* New York: Wiley.

Prokasy, W. F., Jr. (1956). The acquisition of observing responses in the absence of differential external reinforcement. *J. comp. physiol. Psychol.* **49**, 131–4.

Schoenfeld, W. N. (1950). An experimental approach to anxiety, escape, and avoidance behavior. In *Anxiety*, ed. Hoch, P. H. and Zubin, J. New York: Grune and Stratton.

Wehling, H. E. & Prokasy, W. F., Jr. (1962). Role of food deprivation in the acquisition of the observing response. *Psychol. Rep.* **10**, 399–407.

Wyckoff, L. B. (1952). The role of observing responses in discrimination learning: I. *Psychol. Rev.* **59**, 431–42.

(*Manuscript received* 13 *May* 1964)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 285–288
Printed in Great Britain

VARIETY OF BEHAVIOURS AND FREQUENCY OF RESPONDING DURING AVOIDANCE CONDITIONING

By BERNARD LYMAN
*Simon Fraser University, Canada**

Ten female goats were trained in avoidance flexion of the right forelimb. The subjects' qualitative and quantitative activity was recorded. When the first and last tenths of training were compared, there were significant increases in the overall frequency of responding during periods of conditional stimulus impingement and in the overall variety of behaviours, omitting flexion and emotional behaviours, during inter-trial intervals. There were significant decreases in the variety of emotional behaviours during both signal periods and intervals, in the frequency of flexion responses during intervals, and in the frequency of emotional responses during signal periods. The data do not support the assertion that late in avoidance training the organism is quiescent and is mechanically emitting only the conditioned responses to the conditional stimulus.

Some years ago, Culler, Finch, Girden & Brogden (1935) reported that, as training progressed during avoidance conditioning of the dog, irrelevant behaviours dropped out until only finely executed conditioned responses (CR) were made to the conditional stimulus (CS). Although in classical conditioning, with electric shock as the unconditional stimulus (US), emotional responses may remain (Brogden, Lipman & Culler, 1938; Brogden, 1939), Gibson (1952) observed the quieting phenomenon in goats with both classical and instrumental procedures. It appears that the effect is not dependent on use of the Pavlov frame or other restraining devices. When the sheep or goat is conditioned with unrestricted freedom of movement, the animal restricts itself to a small part of the laboratory room and shows little activity other than responding to the CS. Liddell (1950*a*) refers to this phenomenon as 'self imposed restraint' and suggests that the animal's quiescence is more apparent than real (Liddell, 1950*b*). The present paper deals with the question of whether or not the animal is more active at the beginning than at the end of training. The frequency of responding and the variety of behaviours or qualitatively different responses are considered.

Method

The method is described in detail in a previous paper from the same study (Lyman, 1959) and is summarized here.

Subjects. The subjects were ten naïve female goats from the Cornell University Behaviour Farm stock. (Twelve subjects were conditioned, but one developed encephalitis and a second remained underdeveloped three years after completion of the study.) At the start of the experiment, the ages ranged from 15 to 19 months with a mean of 16·3 months.

Procedure. Three subjects were trained at one time. Each subject was restrained lightly near a wall by a stanchion-like apparatus. The CS was a 10 sec buzzer and the US was a mild electric shock. There were fifteen daily trials until the criterion of learning was reached. The criterion was twofold. There had to be at least ten consecutive avoidance responses during a single training session, and the cumulative number of avoidance responses had to be at least twice the cumulative number of unconditioned responses.

* The study was conducted at Cornell University's Behaviour Farm Laboratory.

The observer, using a previously developed check-list, sat in one corner of the room and recorded each gross overt response of each subject on every third trial. The behaviour was recorded as occurring during signal periods (CS impingement) or during interstimulus intervals following those signal periods. Practical limitations precluded records of the individual components of non-unitary responses; e.g., tail flicking refers to a tail flicking episode without regard to the number of discrete flicks of the tail in an episode, and, save for leg flexion, it was not feasible to record the magnitude or duration of the behaviours.

Results and discussion

The mean number of trials to criterion was 243·0 with an S.D. of 90·97.

From the protocol check-lists, each of the observed behaviours was classified as belonging in one of three categories: (*a*) flexion behaviours, (*b*) emotional behaviours, or (*c*) other behaviours.

(*a*) The behaviours classified as flexions were: flexions of the right forelimb, other than unconditioned and avoidance responses, which had an amplitude of at least 2½ in., pawing motions of the right forelimb, and waving which was a pumping movement of the rigidly extended right forelimb. Undonditioned and avoidance flexions were omitted from the data as there was always one or the other for each trial.

(*b*) The behaviours classified as emotional were the more insistent ones of bleating, trembling, and eliminating, and those which could be described as attempts to escape the apparatus—twists head, struggles, and rears.

(*c*) In order of decreasing frequency of occurrence, the remaining observed behaviours were: orienting response, postural adjustment, swallows, grinds teeth, chews cud, nuzzles, flicks tail, flicks ears, snorts, yawns, stretches, coughs, shakes head, attempts to lie down, licks chops, shakes self, sneezes, and scratches.

Trends for each category of behaviour were obtained for each subject by the use of 10 Vincent units based on observed trials from the first trial through the last error trial. The mean number of such trials was 76·6, with S.D. 28·07. The data refer to activity occurring during observed signal periods and interstimulus intervals following those signal periods and are presented in terms of means per subject per Vincent unit.

Table 1. *Mean frequency of responding per subject per Vincent unit during avoidance acquisition* ($n = 10$)

	Behaviour	Vincent units 1	2	3	4	5	6	7	8	9	10
All	Signals	27·5	50·6	56·7	50·0	56·4	57·1	53·8	53·1	53·2	51·3
	Intervals	77·4	92·8	86·8	86·6	84·6	91·4	78·2	83·9	87·0	68·7
Flexion	Signals	3·9	12·4	16·1	11·2	13·1	10·3	6·4	6·3	5·2	5·6
	Intervals	12·4	15·3	11·6	13·2	11·8	14·5	6·6	8·8	3·6	2·6
Emotional	Signals	2·3	3·2	1·7	0·3	0·2	0·2	0·0	0·1	0·2	0·0
	Intervals	7·3	9·7	3·7	1·3	2·2	1·6	0·5	0·9	0·7	0·1
All but flexion and emotion	Signals	21·3	35·0	38·9	38·4	43·0	46·7	47·3	46·7	47·8	45·7
	Intervals	57·7	67·8	71·5	72·1	70·6	75·4	71·0	74·2	82·8	66·0

Table 1 shows the mean frequency of responding; i.e. the mean number of times behaviours or behaviour episodes occurred during a given unit. Table 2 shows the mean variety of behaviours; i.e. the number of qualitatively different behaviours or behaviour episodes without regard to the number of repetitions of a given behaviour

in any unit. These tables show that the organism's overall activity level remains relatively constant throughout training.

Table 3 shows that the overall variety of behaviours, during both signal periods and intervals, is not significantly different at the end of training from what it was at the start. In regard to frequency of responding, there is a significant increase during signal periods and no change during intervals.

Table 2. *Mean variety of behaviours per subject per Vincent unit during avoidance acquisition* ($n = 10$)

		Vincent units									
	Behaviour	1	2	3	4	5	6	7	8	9	10
All	Signals	5·7	5·6	5·2	5·0	5·0	5·6	5·0	5·6	4·8	5·4
	Intervals	8·6	8·5	9·5	10·0	9·3	9·6	9·5	9·7	11·2	9·4
Flexion	Signals	1·1	1·6	1·8	1·7	1·8	2·1	1·7	1·5	1·2	1·3
	Intervals	1·4	1·4	1·8	1·5	1·7	1·5	1·7	1·5	1·1	0·8
Emotional	Signals	1·1	1·5	0·8	0·1	0·2	0·2	0·0	0·1	0·1	0·0
	Intervals	1·5	1·5	1·2	0·8	0·3	0·6	0·3	0·3	0·3	0·1
All but flexion and emotion	Signals	3·5	2·5	2·6	3·2	3·0	3·3	3·3	4·0	3·5	4·1
	Intervals	5·7	5·6	6·5	7·7	7·3	7·5	7·5	7·9	9·8	8·5

Table 3. *Comparison of activity early and late in training* (*Vincent unit* 1 *minus Vincent unit* 10)

		Frequency of responding		Variety of behaviours	
	Behaviour	$\bar{D}$	t	$\bar{D}$	t
All	Signals	−23·8	4·14**	0·3	0·53
	Intervals	8·7	0·86	−0·8	1·01
Flexion	Signals	−1·7	1·21	−0·2	0·69
	Intervals	9·8	3·92**	0·6	2·22
Emotional	Signals	2·3	2·99*	1·1	3·15*
	Intervals	7·2	2·07	1·4	4·12**
All but flexion and emotion	Signals	−24·4	5·19***	−0·6	1·94
	Intervals	−8·3	1·03	−2·8	4·00**

* $P < 0·05$. ** $P < 0·01$. *** $P < 0·001$.

The data are contrary to those reported by Culler *et al.* (1935) and contrary to what sometimes is implied when the activity level is discussed (e.g. Munn, 1961; Underwood, 1949). Although unrecorded changes in duration and magnitude of behaviours may have occurred, the overall activity level, in terms of variety and frequency of behaviours or behaviour episodes, either increased or showed no significant change. The results may differ from Culler *et al.* because of somewhat different procedures (temporal durations differed, and Culler sounded a buzzer upon termination of the CS), of species differences, or of differences in the degree of systematic quantitative observation. Certainly, the significant decreases in both frequency and variety of the insistent flexion and emotional behaviours (Table 3) tend to give the *impression* of a lowered activity level.

The data support Liddell's hypothesis of an apparent rather than real quiescence late in training and have some bearing on habituation hypotheses of learning which assert a dropping out of behaviours. The data suggest that attention should be given to the specific behaviours involved, that a distinction should be made between changes in frequency and changes in variety, and that it should be determined empirically whether such behaviours are relevant or irrelevant, or competing or non-competing ones. In order to determine the meaning or significance of a given behaviour, detailed analyses, including field observations, are necessary.

REFERENCES

BROGDEN, W. J. (1939). The effect of frequency of reinforcement upon the level of conditioning. *J. exp. Psychol.* **24**, 419–31.

BROGDEN, W. J., LIPMAN, E. A. & CULLER, E. (1938). The role of incentive in conditioning and extinction. *Amer. J. Psychol.* **51**, 109–17.

CULLER, E., FINCH, G., GIRDEN, E. & BROGDEN, W. J. (1935). Measurements of acuity by the conditioned-response technique. *J. gen. Psychol.* **12**, 223–7.

GIBSON, E. J. (1952). The role of shock in reinforcement. *J. comp. physiol. Psychol.* **45**, 18–30.

LIDDELL, H. S. (1950*a*). Animal origins of anxiety. In *Feelings and Emotions* (ed. M. L. Reymert). New York: McGraw-Hill.

LIDDELL, H. S. (1950*b*). The role of vigilance in the development of animal neurosis. In *Anxiety* (ed. P. H. Hoch and J. Zubin). New York: Grune and Stratton.

LYMAN, B. (1959). Signal vigilance level and conditioned avoidance acquisition in the goat. *J. comp. physiol. Psychol.* **52**, 89–91.

MUNN, N. L. (1961). *Psychology: the Fundamentals of Human Adjustment*, 4th ed. Boston: Houghton-Mifflin.

UNDERWOOD, B. J. (1949). *Experimental Psychology*. New York: Appleton-Century.

(*Manuscript received 30 December 1964*)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 289–294
Printed in Great Britain

THE EFFECT OF RATIO OF REINFORCEMENT ON PERFORMANCE IN SELECTIVE LEARNING BY CHILDREN

By J. P. DAS,* GEETA SAHU and T. P. PANDA
Utkal University, Bhubaneswar, India

In the present study, children were required to switch out one of two lights on *choice* trials by pressing one of two available keys and to switch out the only light presented on *forced* trials. By appropriate variation of the two stimuli on forced trials, one of the keys was pressed twice as frequently as the other (Expts. I and II) and three times as frequently as the other (Expt. II). The results showed a gradual initial increment of choice for the more frequently reinforced key (Expt. I) followed by a decrement, and the point of maximum difference between the latency of response of the two keys exactly coincided with the highest point in the choice curve (Expt. I). Similar results for the choice behaviour were obtained in the 2:1 and 3:1 groups of Expt. II where both acquisition and reversal were studied. Reversal for the 3:1 group was fast. Results of Expt. I supported Spence's prediction.

The two experiments on children described here attempt to replicate and extend the following observation reported by Spence (1956, pp. 205–14). In a study by Ramond, rats were permitted to choose to press one of two lighted bars on choice trials, and to press the only lighted bar on forced trials. By appropriate manipulation of the ratio of reinforcement, one of the bars (frequent bar) was pressed twice as frequently for a similar amount of reinforcement as the other (infrequent) bar. It was noticed that the choice of frequent bar increased to a certain point and then decreased to a level approximating the number of choices of the infrequent bar. Spence explained this by presuming that the habit strengths for both frequent and infrequent bars build up as trials proceed, reach a maximum point of difference, and then move at a differential rate to a common asymptote. The choice of the frequent bar thus reflected this difference between their respective habit strengths. Logically, the point of maximum difference between the two habit strengths should closely approximate the trial on which the frequent bar was chosen most often. This was empirically borne out in Ramond's data, supporting Spence's hypothesis.

The experimental situation was suitably modified in the present study for testing children. Subjects were required to switch off one of two lights on choice trials and the only light on forced trials. The first experiment reported here closely followed Ramond's study. In the second experiment, an additional 3:1 reinforcement ratio was introduced, and both acquisition and reversal measures were obtained. Reward in the form of candy was given at the beginning and at the end of the experimental sessions, but not on every trial, a procedure which has been successfully adopted in similar experiments with children (Das & Panda, 1963). This was done (*a*) because the subjects did not then associate the candy reward at the end of a trial with 'success', and (*b*) because it was feared that rewarding a choice trial would lead subjects to repeat the specific response on all choice trials.

* This paper was written while the senior author was a visiting professor, Joseph P. Kennedy Jr. Foundation, George Peabody College, Nashville, Tennessee.

Experiment I

Method

Apparatus. The subject sat facing a wooden display panel on which two flash-light bulbs were mounted. These could be switched off by pressing corresponding keys located just below the lights on a horizontal board. At the back of the display panel, the experimenter could switch on any one or both of the lights which started off a chronoscope. Switching off the light stopped the chronoscope. Latency of response was measured by reading the chronoscope to the nearest 10 msec.

Subjects. The subjects were thirty children between 7 and 11 years of age with a mean age of 9 yr. They were taken from the University's elementary school and were divided into two groups matched for age and sex in order to receive the counterbalanced treatments described in the following section.

Design and procedure. Distinguishing the two keys as K_A and K_B, the ratio $K_A:K_B$ was 2:1 for half of the subjects (group I), and for the other half (group II) $K_B:K_A = 2:1$. Subjects received a total of 34 choice and 66 forced trials in blocks of three trials consisting of one choice followed by two forced trials. On choice trials, both K_A and K_B were available, whereas on forced trials the subject had to press the key under the lighted lamp. Availability of K_A or K_B on forced trials was contingent upon a subject's response on choice trials. For group I, if K_A was chosen on a choice trial, the next two forced trials were K_A, K_B or K_B, K_A according to a predetermined random sequence. On the other hand, if K_B were chosen, K_A was presented on the next two forced trials. A corresponding contingency was adopted for group II. Children used in a pilot study had reached their asymptote in reaction latency at the end of 100 trials. On the basis of this, the total number of trials in the final study was fixed at 100.

Subjects were tested individually by one of the authors (G.S). The main test was preceded by ten training trials in which the subject was taught to press the appropriate key to switch off the light. Subjects were instructed to press only one key when both the lights were on. They were further urged to respond as fast as they could. The inter-trial interval was approximately 10 sec.

All subjects were given the Porteus maze test for 11-year level in order to discover differences in performance that might correspond to variation in speed of maze solution.

Results

Four of the thirty subjects persisted in responding to the less reinforced light on choice trials. The results of these subjects have been excluded from further computations.

Table 1

(A) *Percentage of choice of the more frequently reinforced responses for eight blocks of four trials by the 2:1 group (n = 26)*

I	II	III	IV	V	VI	VII	VIII
59·58	57·66	73·04	70·15	70·15	62·46	62·46	53·81

(B) *Percentage of choices made by more and less intelligent subjects (Porteus Maze Test)*

Maze test	n	I	II	III	IV	V	VI	VII	VIII
Above mean	12	56·25	52·00	68·75	68·75	72·75	60·25	64·50	47·75
Below mean	14	62·50	60·75	92·75	71·50	67·50	64·25	60·75	59·00

Frequency of choosing the more frequent response key in choice trials is presented in Table 1. Each trial block consists of four choice trials excluding the 1st and 2nd choice trials. However, the scores on these were not different from those of the first

block. It can be seen from Table 1 (A) that maximum choice occurred on the 3rd block. The general shape of the curve based on these data can be described as gradual rise to a point followed by a decrement to a level lower than the starting-point. This characteristic is very clearly shown in the data reported in Table 1 (B). The above-mean and below-mean groups refer to subjects who had maze solution speeds above and below the mean for the Porteus maze test. The scores reach a maximum of 92·75% choice on the 3rd block and show a consistent decrement.

Table 2 presents mean reaction times for blocks of eight forced trials. The first two forced trials have been excluded, thus leaving 64 trials divisible into eight blocks. Scores have been smoothed by a simple technique of adding the preceding and succeeding scores to the score to be smoothed and dividing the sum by three. However, the first and last scores in the series have been added respectively to their succeeding and preceding scores and their sum has been divided by two. Such a smoothing procedure can be regarded as 'blind' to the experimenters' predictions. It may be observed that for group I, reaction time for the frequent responses seems to have attained stability at the end but not so the reaction time for the infrequent response. However, all the four series of reaction times show progressive decrements characteristic of practice.

Differences between the smoothed reaction times are also presented in Table 2, and have been summed separately for each block. It is apparent that the maximum difference falls in block III. This coincides exactly with the point of maximum rise in frequency of choice (Table 1A). The general agreement between difference in latency and frequency of choice was also very close, the rank order correlation over the eight blocks of trials being $\rho = +0{\cdot}90$.

Table 2. *Smoothed mean reaction time (RT) in msec for frequent and infrequent responses*

Group I. $K_a:K_b = 2:1$	Blocks of eight trials							
	I	II	III	IV	V	VI	VII	VIII
RT (freq.)	711·0	684·3	635·0	596·3	576·0	574·3	571·7	572·5
RT (infreq.)	728·0	707·3	706·3	653·7	640·7	611·3	603·3	591·5
Difference	17·0	23·0	71·3	57·4	64·7	37·0	31·6	19·0
Group II. $K_b:K_a = 2:1$								
RT (freq.)	732·5	708·0	652·7	625·7	596·7	573·3	550·7	540·0
RT (infreq.)	722·5	712·0	696·3	649·0	631·3	604·7	598·0	578·0
Difference	10·0	4·0	43·6	23·3	34·6	31·4	47·3	38·0
Groups I and II. Total difference	27·0	27·0	114·9	80·7	99·3	68·4	78·9	57·0

Supplementary analysis of reaction times for forced and choice trials, as well as for more and less frequently reinforced responses, was carried out. Analysis of variance yielded a significant main effect for each of these variables implying that forced RT was shorter than choice RT ($F = 10{\cdot}05$, $P < 0{\cdot}01$) and that the more frequently reinforced response was speedier than the less frequently reinforced response ($F = 136{\cdot}51$, $P < 0{\cdot}01$). Unexpectedly, choice RT for the frequent response was longer than choice RT for the infrequent response. This resulted in a significant interaction between the main effects ($F = 10{\cdot}1$, $P < 0{\cdot}01$).

EXPERIMENT II

Two things were different in this experiment. (*a*) The reinforcement ratio was varied. Two groups of subjects received a 2:1 ratio of reinforcement and another two received a 3:1 ratio. (*b*) Reversal, in addition to acquisition, was included in the test. On reversal trials, the less frequently reinforced response acquired the reinforcement ratio of the more frequently reinforced response, and vice versa. This experiment also differed from the previous one in the massing of trials. The inter-trial interval was reduced to 5 sec., half of the interval in Expt. I.

Method

Subjects were eighty children taken from the elementary schools of the area adjacent to the campus. Their age structure was the same as those of the previous sample—7 to 11 years old with a mean around 9.

All subjects were tested individually by one of the authors (T.P). They were tested in an empty office or classroom of the school. Owing to some practical difficulties, the chronoscope could not be used in these schools, hence only frequency of choice was recorded. Consistently with the previous experiment, candy was distributed before and after the test.

There were two counterbalanced groups of twenty subjects each for the 2:1 reinforcement ratio treatment ($K_A:K_B = 2:1$; $K_B:K_A = 2:1$), and two similar groups for the 3:1 reinforcement ratio ($K_A:K_B = 3:1$; $K_B:K_A = 3:1$).

In the 3:1 group, blocks of four trials starting with a choice and followed by three forced trials were given. The choice of the frequent response led to two frequent responses and one infrequent response on the forced trials in a predetermined random order, whereas the choice of the infrequent response was followed by the frequent response on all three forced trials.

Both the groups received 34 choice trials in acquisition and 24 in reversal. The number of forced trials corresponded to the ratio of reinforcement for each group.

Table 3. *Percentage of choices of the more frequently reinforced response for eight blocks of acquisition trials and six blocks of reversal trials (four trials/block)*

(A) Group I. 2:1 ($n = 40$)

	I	II	III	IV	V	VI	VII	VIII
Acquisition	46·9	57·5	62·5	54·4	52·5	46·9	56·2	76·2
Reversal	77·5	47·5	36·2	41·9	39·4	15·0	—	—

(B) Group II. 3:1 ($n = 40$)

	I	II	III	IV	V	VI	VII	VIII
Acquisition	50·0	50·6	60·6	78·1	71·2	70·0	68·7	86·9
Reversal	83·7	44·4	35·6	19·4	11·9	5·6	—	—

Results

Percentages of choice of the more frequently reinforced response are given in Table 3. The 2:1 group (Table 3A) has an initial choice of 46·87 gradually increasing to 62·50 and dropping to a level of 46·87 again on the 6th block. But, unlike the data of Expt. I, the choice subsequently increases and reaches a very high percentage on the last trial block. This double-crested character is individually found in each of the two counterbalanced groups whose scores have been pooled to prepare Table 3A. Reversal choices, likewise, consistently decrease until on the 4th block the percentage choice of the now infrequent response rises by 5%, but subsequently decreases to a very low level in the last block (Fig. 1).

The acquisition scores are consistent with those previously obtained if one considers the curve between the 1st and the 6th block. This may be viewed as another

replication study supporting Spence's hypothesis. However, the subsequent rise in choice following the 6th block of trials cannot be easily explained in the light of this hypothesis. It can best be attributed to massing which appears to increase the preference for the more frequently reinforced response.

Reversal for the 2:1 group did not lead to a gradual decrement in choice of the now infrequent response. The choice decreased to about 36% in the 3rd block, but then increased for two blocks before decreasing to 15%. If one conceives of the reversal process as a mirror image of acquisition, this moderate rise in choice during reversal may be in line with Spence's prediction for acquisition (Fig. 1).

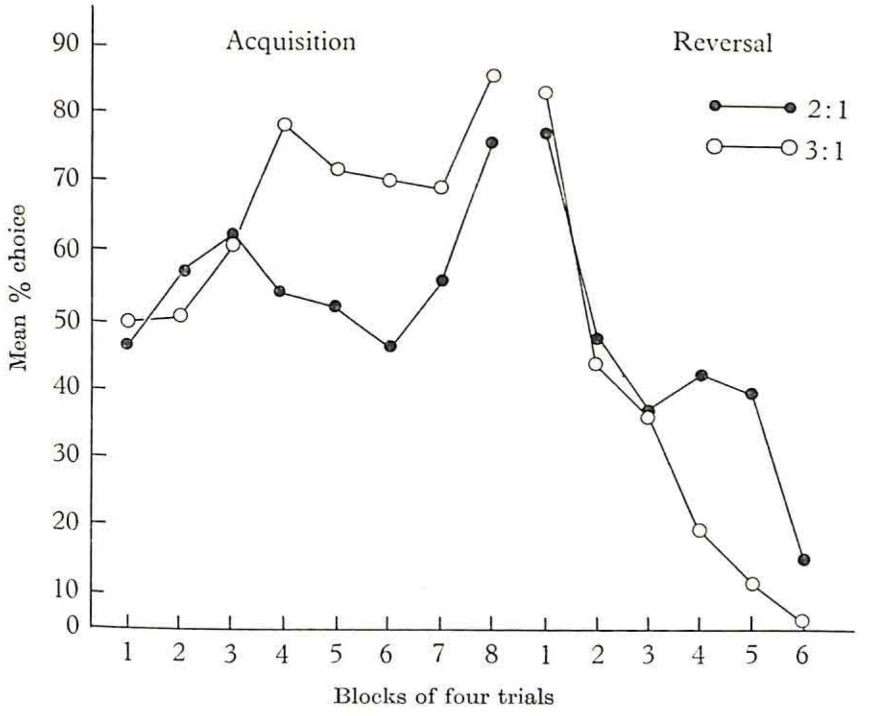


Fig. 1. Percentage of choices during acquisition and reversal as influenced by two ratios of reinforcement ($n = 40$ in each group).

Choice scores for the 3:1 group are given in Table 3B. The choice of the frequent response reached a maximum on the 4th block followed by a slow decrement. But in the last trial block, the choice increased again. Analysis of the choice scores for the two counterbalanced groups separately reveals that in one the choice increased on the 8th block for the second time, whereas in the other group it did so from the 6th block. Reversal scores, however, show a gradual and consistent decrement following an initial sudden decrement of about 44% from the 1st to the 2nd block. Choice of the now infrequent response becomes less than 6% in the last block of reversal trials. This is much lower (9%) than the comparable figure for the 2:1 group.

In acquisition, the choice score for the 3:1 ratio does not return to anywhere near the 50% level, whereas it fell to about 47% in the 2:1 group. This cannot be attributed to sampling differences between two groups of subjects; like the 2:1 group, the 3:1 group also had an initial chance score on the first block of trials.

DISCUSSION

In Expt. I, it was found that Ramond's experiment on rats could be replicated with children. Prediction of choice behaviour from differential growth of habit strengths for the frequent and infrequent response was also achieved. However, the latency data were not very regular although the coincidence of latency difference and frequency of choice was evident even in the unsmoothed data. In addition to this, the overall correspondence between latency and frequency is convincingly shown by a correlation of $+0{\cdot}90$ between them.

The shape of the choice curve was partially replicated in the 2:1 group of Expt. II. Its subsequent rise in the choice of the frequent response in the 2:1 as well as in the 3:1 groups needs to be explained. Can this be done within the theoretical framework which explains the choice data in the first experiment? Considering the 3:1 group, because it is most dissimilar in its choice compared with the 2:1 group of Expt. I, does the choice behaviour reflect differential habit growths? In this group the frequent habit receives three times as much reinforcement as the less frequent one. As a result, it may be inferred that the 'frequent' habit grows much faster than the 'infrequent' one, and that the two habit strengths cannot come as close as they would in the 2:1 group. But if Spence's hypothesis still holds good for a 3:1 reinforcement ratio, the *pattern* of difference between the two habits would be the same as that observed in the 2:1 condition. This is indeed supported by the data (Table 3B). Habit strengths for such simple key pressing responses may be conceived to reach a temporary asymptote in human subjects. This would be characteristic of a plateau in performance curves. Following a plateau stage, the two habits may start to diverge further, which may account for the rise in the choice curve for the second time. Would this subsequent divergence lead to a similar phase of maximum difference preceding a gradual reduction of their difference as the habit strengths approach another asymptote? This seems unlikely. On the other hand, the results of Expt. II appear to suggest that behaviour at this stage cannot be predicted from Spence's theory. For one thing, habit strength cannot be measured by response latency which would have reached an upper limit at the 8th block of trials. Thus, although the first stage of performance can be adequately explained by the differential build up of habit strength commensurate with the ratio of reinforcement, the second shift in performance needs a new explanatory construct. It may be suggested that as trials are continued beyond the typical Spencian effect, response to the frequent key begins to acquire the characteristics of a conditioned response. That is, every choice trial becomes a conditioned stimulus evoking choice of the frequent key as its CR. This CS–CR bond would grow in strength and lead ultimately to a 100% choice of the frequent key. Such a limiting condition is already embodied in the 87% terminal choice of the frequent key by the 3:1 group.

REFERENCES

DAS, J. P. & PANDA, K. C. (1963). Two-choice learning in children and adolescents. *J. gen. Psychol.* 68, 203–11.

SPENCE, K. W. (1956). *Behavior Theory and Conditioning*. New Haven: Yale University Press.

(*Manuscript received* 16 *June* 1964)

Brit. J. Psychol. (1965), **56**, 2 and 3, *pp.* 295–302
Printed in Great Britain

THE DIFFICULTY AND GRADED SCORING OF ELITHORN'S PERCEPTUAL MAZE TEST

By ANN D. M. DAVIES

Medical Research Council, Unit for Research on Occupational Aspects of Ageing, University of Liverpool

AND M. G. DAVIES

Department of Physics, University of Liverpool

This paper describes an attempt to relate the subjective difficulty of a maze test, as provided by the success of a panel of subjects in solving it, with certain definable physical parameters of the maze. By assuming that the fewer solution paths there are through the maze the harder it is to find one, the correlation between subjective and calculated difficulty is $+0{\cdot}77$. By empirically including the complications of maze saturation and of the number of 'legs' in the solution path, this value is increased to $+0{\cdot}94$. Using these ideas, a weighted scoring scheme is suggested whose values correlate more highly with subjects' Progressive Matrices scores than does the right-wrong value.

Elithorn and his colleagues have developed a number of versions of a perceptual maze test which has been found sensitive to brain damage (Elithorn, 1955; Elithorn, Kerr & Mott, 1960; Elithorn, Kerr & Jones, 1963; Elithorn, Jones, Kerr & Lee, 1964). The test can, however, be regarded as some sort of measure of intelligence (Elithorn *et al.* 1963) and as such was included among a battery of tests given to a large number of adult subjects of all ages and occupations, in the Liverpool area.

An example of the test item is shown in Fig. 1. The maze takes the form of an inverted triangle, and starting from the apex, the subject has to trace a path in an upward direction, keeping to the lines and passing through as many dots as he can—three in Fig. 1—terminating his path on the base of the triangle. The maze items can be conveniently described by their size or rank R (the number of rows, not including the apex) and by the number of dots distributed over the lattice points (or the ratio of junctions filled to the total number of lattice points – the saturation).

The test series consisted of six mazes of rank 12, six of rank 16 and six of rank 20. Each group contained two mazes each of 20, 30 and 40 % saturation. The rank 12 mazes, the easiest ones, were presented first, then the rank 16 and finally the rank 20 items. As administered in the present study, subjects were first told the rules and allowed to work through three demonstration mazes, one of which was Fig. 1. In each case, they were told the number of dots they must find on their solution path. An additional maze was included to introduce the subject to the idea of working without knowledge of the maximum. Subjects were then presented with a series of eighteen mazes in the 'without knowledge' condition and allowed 12 min for completion. Testing was done individually. Subjects who completed all eighteen in less than 12 min were encouraged to try to improve their performance.

Although the literature describes various methods of presentation of the Perceptual Maze Test, the only method of scoring, to date, has been to award the subject a point if he traces a path through the maximum possible number of dots; if he fails to do this, he scores nothing on that maze. This scheme has two shortcomings: it makes no distinction between an easy and a difficult maze and it makes no distinction

between a poor solution and one which is good but not complete. This paper is concerned with a scoring technique which obviates both these disadvantages. It is based on the idea that there are a large number of distinguishable paths through the maze, any one of which might be followed, making a purely random choice of direction at each lattice point. The 'difficulty' of the maze is related to the number of paths through the maximum number of dots ($\hat{m}$) on the solution path, and the subject score is related to the difficulty of finding a path through the m dots he has attained.

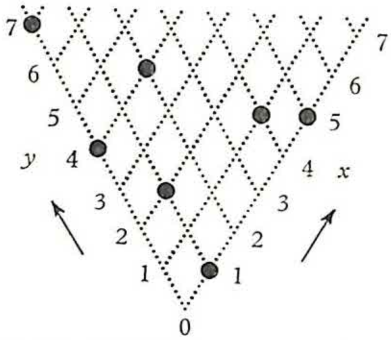


Fig. 1. One of the demonstration mazes, $R = 7$, $\hat{m} = 3$.

The calculated difficulty

The lattice in Fig. 1 can be described by a system of Cartesian co-ordinates (x, y). From the apex $(0, 0)$ we may move to $(1, 0)$ where there is a dot or to $(0, 1)$ where there is none. To reach $(2, 0)$ we must pass through $(1, 0)$ and therefore through a dot. To reach $(1, 1)$ we may pass through $(1, 0)$ with its dot or through $(0, 1)$ and so no dot. To reach $(0, 2)$ we pass through no dots.

If u_m is the number of separate paths up to a lattice point which pass through m dots, we have for the second rank:

$$\begin{array}{lll} (2, 0) & u_0 = 0 & u_1 = 1 \\ (1, 1) & u_0 = 1 & u_1 = 1 \\ (0, 2) & u_0 = 1 & u_1 = 0 \end{array}$$

A simple inspection of the four points on the third rank shows that:

$$\begin{array}{lllll} (3, 0) & u_0 = 0 & u_1 = 1 & u_2 = 0 & \Sigma u_m = 1 = \binom{3}{0} \\ (2, 1) & u_0 = 1 & u_1 = 2 & u_2 = 0 & \Sigma u_m = 3 = \binom{3}{1} \\ (1, 2) & u_0 = 0 & u_1 = 2 & u_2 = 1 & \Sigma u_m = 3 = \binom{3}{2} \\ (0, 3) & u_0 = 1 & u_1 = 0 & u_2 = 0 & \Sigma u_m = 1 = \binom{3}{3} \end{array}$$

The number of paths to a given point is one of the binomial coefficients forming Pascal's triangle. For the lattice points on the $(x+y)$th row,

$$\sum_{m=0}^{m} u_m = \binom{x+y}{x} = \binom{x+y}{y}.$$

This procedure can be continued for the entire maze and the separate values of $u_0, u_1, \ldots$ at each point on the last row can then be summed. Thus

$$U_m = \sum_{x=0}^{R} u_m.$$

The relation between the distribution of dots over the lattice and values of U_m is discussed elsewhere (Davies & Davies, 1965).

The values of U_m for Fig. 1 are shown in Table 1.

Table 1. *Values of U_m for the maze item shown in Fig. 1*

m	0	1	2	3
U_m	14	59	49	6

It will be seen that

$$\sum_{m=0}^{3} U_m = 128 = 2^7.$$

In general,

$$\sum_{m=0}^{\hat{m}} U_m = 2^R.$$

The maximum number of dots through which it is possible and permissible to pass is 3 in Fig. 1 and there are six distinct paths through them. The probability of following one by chance is 6 in 128 or $U_{\hat{m}}/2^R$. Clearly, the fewer the number of paths there are through $\hat{m}$ dots, the more difficult it is to find one of them. It seems reasonable to take this measure, or rather its inverse, as a measure of the calculated difficulty of the maze. (It should be remarked that this would not be true if the dots were arranged to form some clear chain from apex to base. The dots however were fitted on to the lattice in a random manner and the probability of forming such regular patterns is low.) The numbers obtained by evaluating $2^R/U_{\hat{m}}$ are very large, and it is convenient to take their logarithm to the base 10. So we write for the calculated difficulty of the maze

$$D_c = \log_{10}(2^R/U_{\hat{m}}).$$

There are two comments to be made about this score. First, the number of paths varies exponentially with R, and a logarithmic function is therefore a reasonable one to use to obtain workable figures; secondly, if we were to add to the maze a further unfilled row of lattice points, the value of the total number of paths through the maze would double, but so would each U_m; the ratio of 2^R (total number of paths) to $U_{\hat{m}}$ (the paths through the maximum number of dots) would thus remain unaltered. This is a property which is clearly necessary for any scoring scheme.

The value of D_c was calculated for each maze in the set of eighteen that was used.

The empirical difficulty of the mazes

Empirically, an easy maze is one which a high proportion of the subjects solve correctly. Confronted by a difficult maze a small fraction of the population will find a path through $\hat{m}$ dots and other fractions through $\hat{m}-1$, $\hat{m}-2$, and so on. Table 2 shows the performance of the panel on a maze of 40 % saturation and rank 20.

Table 2. *The performance of the panel on a maze of rank 20, 40 % saturation*

		16	15	14	13	12	11	10	9	8	7	6	Total
No. of dots traversed (m)													
No. of subjects (n):	Men	14	27	33	25	12	3	1	2	—	—	1	118
	Women	8	16	32	22	6	7	3	1	1	—	2	98

The mean value of m for the population as a whole is

$$m_g = \frac{\Sigma nm}{\Sigma n}.$$

The value by which m_g falls short of $\hat{m}$ is taken to be the empirical difficulty of the maze.

$$D_e = \hat{m} - m_g$$

(2·2 for the men, 2·6 for women on the maze referred to in Table 2). This was evaluated for each of the eighteen mazes using a panel of the 118 men and 98 women who had completed all eighteen mazes. (The total panel consisted of 300 men and 240 women uniformly spread between 20 and 79 years of age and in occupations representative of the national distribution.) D_e was evaluated separately for men and women because men consistently find less difficulty in solving the mazes than women (Davies, 1965). However, the rank order of difficulty of the mazes was very similar for men and for women (Spearman's $\rho = +0{\cdot}97$) and for subsequent analysis the mean of the ranks for men and women was used.

Comparison of calculated and empirical difficulties

The mazes could now be arranged in rank order of calculated difficulty and also in rank order of empirical difficulty. The correlation between the two sets of ranks was computed and a significant relationship between the ranks was found ($\rho = +0{\cdot}77$, $P < 0{\cdot}01$). The relation between empirical and calculated difficulty could be made closer in two ways. Consider two mazes, M_1 and M_2 of the same rank, such that M_2 contains M_1, together with additional dots which do not alter the value of $\hat{m}$. In these circumstances the effect of the extra dots will largely be to obscure the best path. (They might of course increase $U_{\hat{m}}$ without increasing $\hat{m}$). The saturation of the maze (the ratio of number of dots to number of lattice points) may therefore be expected to have some effect on the calculated difficulty.

It was possible to examine the effect of both size and saturation on empirical difficulty by an analysis of variance. The difficulties of mazes having the same rank and saturations were averaged for men and women together and these average empirical difficulties are shown in Table 3. The analysis of variance revealed that both the size of the maze and its saturation influenced empirical difficulty (for size $F = 62{\cdot}5$, $P < 0{\cdot}001$; for saturation $F = 7{\cdot}8$, $P < 0{\cdot}05$). It will be seen that within the range of size and saturation explored size has a much greater influence on empirical difficulty than has saturation. These results are in accord with those of Elithorn *et al.* (1964).

Table 3. *Empirical difficulty of mazes classified by rank and saturation*

	Rank R		
Saturation (%)	12	16	20
20	0·28	1·10	1·65
30	0·40	1·25	2·22
40	0·80	1·52	2·55

In introducing saturation into the expression for calculated difficulty it seems reasonable from the above analysis for the calculated difficulty to increase with saturation, but not very rapidly. Thus

$$D'_c = \log_{10}(2^R . s^a / U_{\hat{m}}).$$

Using various integral values of a, values of D'_c were computed and the corresponding rank orders compared with the rank order of empirical difficulty. The closest correspondence was obtained with $a = 4$, when the correlation coefficient became 0·87.

A second 'improvement' was effected by arguing that, other things being equal, a maze would be found more difficult the greater the number of separate straight line portions (or 'legs') that were required to pass through $\hat{m}$ dots. It is possible to complete Fig. 1 in three legs. The minimum number of legs (denoted by l) to pass through $\hat{m}$ dots was found by inspection for each maze. Again, it seemed that difficulty should not vary very rapidly with l and l was introduced into D_c as l^b where b is a second empirical constant. Thus

$$D''_c = \log_{10}(2^R . s^4 . l^b / U_{\hat{m}}).$$

Using various values of b the rank orders of D''_c were compared with the rank order of D_e when the correlation coefficient became $+0·94$, with $b = 4$ or 5.

The subject score

The values of U_m for each value of m for one of the rank 16 mazes with 30% saturation is shown in Table 4.

Table 4. *Values of U_m for a maze where $R = 16$, $s = 0·3$, $\hat{m} = 8$*

m	0	1	2	3	4	5	6	7	8
L_m	2,214	10,611	20,078	19,366	9,716	2,831	617	97	6

If a subject made a random choice at each lattice point it will be seen that there is a high probability that he would pass through 2 or 3 dots, and there is a relatively low probability of his passing through 6, 7 or 8 dots by following a random path. If a subject manages to find a path through 7 dots one may be fairly certain that he has made a conscious departure from a random performance, and is presumably exercising his intellectual and perceptual resources to do so. By analogy with the last section a measure to express the performance of a subject finding a path linking m dots might be

$$S_1 = \log_{10}(2^R / U_m).$$

Alternatively, one may argue that it is the deviation from chance level performance which should be rewarded. In this case

$$S_2 = \log_{10}(U_{\tilde{m}} / U_m),$$

where $U_{\tilde{m}}$ denotes the number of paths through the most frequently occurring value of m—20,078 in the above table. S_2 has the advantage over S_1 that the chance level is not rewarded at all. Both S_1 and S_2 display the defect that if a subject should traverse less than $\hat{m}$ his score would increase again; the instructions were, however, to find the path going through the maximum number of dots and none of our subjects showed such perverse ingenuity.

Again, as in the last section, we may argue that a subject's score should increase a little with the saturation of the maze. This generates two further scores

$$S_3 = \log_{10}(2^R . s^4 / U_m) \quad \text{and} \quad S_4 = \log_{10}(U_{\tilde{m}} . s^4 / U_m).$$

The argument for including saturation in the scoring system is not too convincing since although extra dots might not affect $\hat{m}$ they would certainly affect other values

of $U_{m'}$ and U_m has already appeared in the formula. There is, incidentally, no point at all in using a score which takes into account the number of legs used, since this clearly only applies when $\hat{m}$ is achieved.

The values of S_1, S_2, S_3 and S_4 for each subject on each maze were computed and summed for each subject. Thus the complete weighted score for each subject is given by $T_1 = \sum^{18} S_1$, and similarly for T_2 T_3 and T_4.

Comparison of subject scores and non verbal intelligence

All our subjects had previously attempted Raven's Progressive Matrices test (1938 version). A comparison between the number of Matrices solved and T_1, T_2, T_3 and T_4 was therefore possible. The value T_0—the number of completely solved mazes was also included. The product-moment correlations (men and women combined) between T_0 and Progressive Matrices score was $+0{\cdot}53$, that between T_1 and Matrices score $+0{\cdot}61$. This increase in the correlation coefficient is statistically significant ($t = 3{\cdot}67$ $P < 0{\cdot}01$). The coefficients for men and women separately are shown in Table 5.

It will be seen that there is virtually nothing to choose between the four measures T_1, T_2, T_3 and T_4. The performance of women subjects correlates slightly less with Matrices score than does that of men. There is, of course, a marked difference in absolute performance, as mentioned above. The significant difference in the correlation based on the right-wrong score and that based on the weighted score shows that the weighted score, in accordance with the theory, is a more sensitive measure of maze performance. It may be mentioned, in passing, that the Progressive Matrices score is itself based on a right-wrong analysis and suffers from both the defects mentioned in the Introduction above. The difficulty of the Matrices rises very sharply. That of the maze test also increases in order of presentation and one might reasonably expect a higher correlation between the two unweighted scores than between one unweighted and one weighted score. The difference between the relation of T_0 and T_1 etc. to other tests may therefore be greater than appears at first sight.

The great advantage of Elithorn's mazes is that their structure permits a test which may be continuously variable from very easy to very difficult. The structure of the mazes also allows a scoring scheme to be devised which is based on clearly definable quantities and assumptions, whether those made above, or others. These advantages make the mazes merit consideration as intelligence test items, though the test has not been directly validated with this object in mind. Owing to the primitive scoring of other test items, no fully satisfactory test of a weighted score would appear to be available at present.

Table 5. *Product-moment correlations between various measures of maze performance and score on the Progressive Matrices test* (1938)

	T_0	T_1	T_2	T_3	T_4
Men	$+0{\cdot}54$	$+0{\cdot}62$	$+0{\cdot}62$	$+0{\cdot}62$	$+0{\cdot}61$
Women	$+0{\cdot}52$	$+0{\cdot}60$	$+0{\cdot}59$	$+0{\cdot}58$	$+0{\cdot}60$

Finally, Table 6 gives values of $S_2 = \log\,(U_{\tilde{m}}/U_m)$ for each of the eighteen mazes. It is only applicable when the subject does not know the value of $\hat{m}$. A path linking less than $\tilde{m}$ dots is given no credit and is denoted by '—' in the table. To use the

Table 6. *Conversion of raw scores to weighted scores for the perceptual maze test (version V. 0.1)*

Order of presentation	Code no.	Rank	Saturation (%)	No. of dots linked 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	122	12	20	0·00	0·13	0·50	1·16	1·76	2·48	—	—	—	—	—	—	—	—	—
2	106	12	30	—	0·00	0·25	0·74	1·28	1·93	—	—	—	—	—	—	—	—	—
3	103	12	40	—	—	0·00	0·22	0·72	1·65	2·39	—	—	—	—	—	—	—	—
4	109	12	40	—	—	—	0·00	0·22	0·75	1·54	2·54	—	—	—	—	—	—	—
5	107	12	30	—	0·00	0·06	0·34	0·95	2·62	—	—	—	—	—	—	—	—	—
6	133	12	20	—	0·00	0·20	0·66	1·48	—	—	—	—	—	—	—	—	—	—
7	101	16	40	—	—	—	—	—	0·00	0·15	0·52	1·16	2·17	—	—	—	—	—
8	110	16	30	—	—	0·00	0·13	0·44	0·94	1·79	—	—	—	—	—	—	—	—
9	105	16	20	0·00	0·01	0·24	0·52	0·88	1·23	1·79	2·74	—	—	—	—	—	—	—
10	129	16	20	0·00	0·02	0·32	0·85	1·51	2·32	3·52	—	—	—	—	—	—	—	—
11	113	16	30	—	—	—	0·00	0·03	0·20	0·52	1·00	1·67	2·69	—	—	—	—	—
12	134	16	40	—	—	—	—	—	—	0·00	0·22	0·75	1·77	3·40	—	—	—	—
13	111	20	20	—	—	—	0·00	0·08	0·37	0·82	1·45	2·48	5·12	—	—	—	—	—
14	108	20	30	—	—	—	—	0·00	0·01	0·18	0·50	0·98	1·66	2·53	3·96	—	—	—
15	104	20	40	—	—	—	—	—	—	—	0·00	0·07	0·23	0·51	1·00	1·80	3·29	—
16	120	20	40	—	—	—	—	—	—	0·00	0·05	0·18	0·40	0·72	1·16	1·79	2·67	3·87
17	102	20	30	—	—	—	—	—	0·00	0·12	0·46	0·98	1·74	2·90	4·99	—	—	—
18	114	20	20	—	—	0·00	0·06	0·27	0·61	1·12	1·79	2·61	3·54	4·47	—	—	—	—

scoring scheme, count the dots (m) linked by the subject on a maze and obtain the weighted score for that maze by locating the appropriate row and column; e.g. if 6 dots are linked in maze 103, the subject is awarded 0·72 for this maze. The subjects total score, T_2, is the sum of all credits for the number of mazes attempted, hence the maximum weighted score = 55·45.

The writers would like to acknowledge the assistance of Mr J. M. Watt, Department of Numerical Analysis, University of Liverpool, who wrote the programme giving U_m for each value of m on any maze; also the computational assistance given by Miss P. A. Robinson of the Medical Research Council Unit for Research on Occupational Aspects of Ageing. The writers are also indebted to Dr A. Elithorn who supplied the test material.

REFERENCES

DAVIES, A. D. M. (1965). The perceptual maze test in a normal population. *Percept. mot. Skills*, **20**, 287–93.

DAVIES, M. C. & DAVIES, A. D. M. (1965). Some analytical properties of Elithorn's Perceptual Maze. *J. math. Psychol.* **2**, 371–80.

ELITHORN, A. (1955). A preliminary report on a perceptual maze test sensitive to brain damge. *J. neurol. neurosurg. Psychiat.* **18**, 287–92.

ELITHORN, A., KERR, M. & MOTT, J. (1960). A group version of a perceptual maze test. *Brit. J. Psychol.* **51**, 19–26.

ELITHORN, A., KERR, M. & JONES, D. (1963). A binary perceptual maze. *Amer. J. Psychol.* **76**, 506–8.

ELITHORN, A., JONES, D., KERR, M. & LEE, D. (1964). The effects of the variation of two physical parameters on empirical difficulty in a perceptual maze test. *Brit. J. Psychol.* **55**, 31–37.

(*Manuscript received* 3 *August* 1964)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 303–304
Printed in Great Britain

RESPONSE-FACILITATION ON REPETITION OF A LIMB MOVEMENT

By PATRICK RABBITT*
Section on Ageing, N.I.M.H. National Institutes of Health, Bethesda, U.S.A.

In serial choice-response tasks responses are made faster when they are repeated than when they follow other responses. This 'response-repetition effect' was first observed under conditions in which a unique response was made to each member of a set of different signals (Hyman, 1953; Bertelson, 1961, 1963). Repetition implied the repetition of a signal, and possibly, therefore, of the entire sequence of events involved in its discrimination and in the selection of an appropriate response. The data begged the question of whether the repetition of all components in this process contribute equally to the observed reduction in response latency.

Bertelson (1965) has since examined a task in which the appearance of either of two signals called for one response, and the appearance of either of another two required a second response. Repetition effects were still obtained when responses were repeated even though signals were not. In short, repetition of a signal is not a prerequisite for a repetition effect. The present experiment was carried out to determine whether a repetition effect might be obtained when neither a signal nor a precise pattern of muscular movements is repeated.

Method

Eighteen young right-handed people, aged 18–29 years, none of whom showed any insight into the purpose of the experiment, were tested with the Birren (1962) 'Psychomet'. Ten signal lights were mounted 1 in. apart in a horizontal row across the upper half of the subject's control panel. Each light was set 6 in. vertically above a 1 in. square switch closed by any light contact. The ten signal lights and associated switches may be imagined to be numbered from 1 to 10 reading from left to right across the control panel. The subject responded to the onset of any light 1 through 5 by touching the switch vertically beneath it with the forefinger of his left hand, and correspondingly used the forefinger of his right hand to answer lights 6–10. A correct response switched off the signal light and another light in a pre-programmed sequence came on within 20 msec. Programmed sequences of 300 signals and responses were based on tables of random decimal digits with the added constraints that each signal light appeared equally often and no signal ever followed itself. Thus the probability that a response with one hand would follow a response with the other hand was slightly greater ($P = 0{\cdot}556$) than the probability that two successive responses would be made with the same hand ($P = 0{\cdot}444$).

Each subject was practised for twenty to thirty responses until it was clear that he understood the task. The program was then changed and he ran through a sequence of 300 signals and responses under instructions to respond as quickly and accurately as possible. Intervals between successive responses were timed to within 0·01 sec, and recorded by a SETAR (Welford, 1952), which also identified each signal presented and the response made to it; a print-out of the record was obtained and scanned by eye to determine responses following responses made with the same hand, and responses following responses made with the opposite hand. Mean response latencies in each of these categories were calculated for each subject separately to three significant figures. The means and standard deviations set out in Table 1 are derived from the mean response times

* Now returned to Medical Research Council Applied Psychology Research Unit, Cambridge, England.

obtained for each individual subject. Latencies of errors and of error-correcting responses are known to be atypical and so were excluded from this computation. No subject made more than five errors during a run.

Table 1. *Latencies of responses following responses made with the same hand compared with latencies of responses following responses made with the opposite hand*

(Means of means of eighteen subjects)

	Latencies (msec.)	
	Mean	S.D.
Responses following responses with the same hand	534	38
Responses following responses with the contralateral hand	581	42

Results and discussion

Responses following responses made with the same hand were significantly faster than responses following responses made with the opposite hand ($t = 3{\cdot}58; P < 0{\cdot}01$). The probability of alternations between hands was greater than the probability of repetitions of the same hand, so that this effect cannot be due to expectations based on learning of the transition-structure of the sequences of signals presented.

These data show that a 'response-repetition effect' may be obtained when neither a particular signal nor a particular sequence of muscle movements is repeated. A sufficient common feature appears to be the selection of the limb with which the response is made. Bertelson's (1963) results imply that this effect cannot be related to transient improvements in muscular tone in a limb which has recently been used. Thus the effect in this case must depend upon the repetition of only one of a series of decisions in the central nervous system which may collectively be called the 'program' for the selection and organization of a response.

References

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quart. J. exp. Psychol.* **12**, 90–102.

Bertelson, P. (1963). S–R relationships and reaction-times to new versus repeated signals in a serial task. *J. exp. Psychol.* **65**, 478–84.

Bertelson, P. (1965). Serial choice reaction-time as a function of response vs signal-and-response repetition. *Nature, Lond.* (in the Press).

Birren, J. E., Riegel, K. F. & Morrison, D. F. (1962). Age differences in response-speed as a function of controlled variations of stimulus conditions. Evidence of a general speed factor. *Gerontologia*, **6**, 1–18.

Hyman, R. (1953). Stimulus information as a determinant of reaction-time. *J. exp. Psychol.* **45**, 188–96.

Welford, N. (1952). An electronic digital recording machine—the SETAR. *J. scient. Inst.* **29**, 1–14.

(*Mannscript received* 21 *August* 1964)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 305–307
Printed in Great Britain

A NOTE ON SOME CRITICISMS OF THE MOWRER/EYSENCK CONDITIONING THEORY OF CONSCIENCE

By H. J. EYSENCK
Institute of Psychiatry, University of London

In a recent paper Argyle (1964) has presented a brief criticism of the Mowrer/Eysenck theory which attempts to account for the process of socialization and the development of conscience in terms of Pavlovian conditioning (Eysenck, 1964). Argyle makes several criticisms which, however, seem to be based on a misunderstanding of the theory. For instance, he argues that according to this theory: 'The more children are punished the stronger should be the superego. As far as physical punishment is concerned the opposite is found.' This argument confuses the distinction between punishment and aversive conditioning, and as this confusion appears quite widespread it seems desirable to put it right. The Mowrer/Eysenck theory makes use of the concept of conditioning in which the time interval between CS and UCS is well known to be extremely important; aversive conditioning only takes place if the UCS is administered very soon after the occurrence of the CS. Punishment is an all-embracing term which would include aversive conditioning but also the administration of UCS at time intervals far exceeding those at which it could reasonably be expected to aid in the formation of conditioned responses. Thus the occurrence of punishment as such may be, and often is, quite irrelevant to the occurrence of conditioning. I have discussed this point at great length in *Crime and Personality* because it seems to me a cogent argument against the hypothesis that legal punishment is efficacious in producing control over antisocial behaviour in situations where the individual concerned is not under surveillance. My position is thus identical with that of Argyle when he says that: 'The role of physical punishment seems to be restricted to control of behaviour in the presence of the punitive agent, but has little carry-over to other situations.' Thus Argyle advances the same criticism of legal punishment but seems to regard it as a criticism of conditioning theories. This does not seem to be logical or permissible. The type of conditioning I had in mind was that which occurs very early in the child's life when a misdemeanour is *immediately* followed by a slap, withdrawal of love, or some other punishment administered within the time interval known to produce Pavlovian conditioning. (It might be added in parenthesis that the fact that 'numerous studies show that delinquents have received more physical punishment than others' may easily be explicable in terms of my theory by suggesting that it is in these children that early conditioning has failed and that it is because of this failure that punishment was later administered.)

Argyle did not consider another point to which I have devoted a chapter. My argument is, and I have supported it with some independent evidence, that emotion may lead to the consolidation rather than the suppression of criminal behaviour, and that punishment may, through increasing motivation, have the effect of promoting rather than suppressing the behaviour which is being punished. There is ample evidence for this proposition (Church, 1963) and taken in conjunction with the points

made in the preceding paragraph these considerations seem to render nugatory Argyle's earlier criticism.

Argyle goes on to argue that 'It is withdrawal of love which seems to produce true superego formation; if this can be interpreted as a more severe kind of punishment than physical punishment the theory could stand, but there is nothing within the theory to suggest why this should be so.' This is a complicated point but it may perhaps be argued that the comparatively light punishments which are meted out to most children by their parents acquire a stronger negative reinforcement value because they are regarded as evidence of withdrawal of love in those children where there is a strong positive relationship within the family. Evidence for this comes, although somewhat indirectly, from the Solomon experiment described in detail in *Crime and Personality*; he found that aversive conditioning in puppies was much more effective in those who had been fed by the experimenter (who administered the aversive treatment) than in those who had been fed automatically.

Argyle concludes his brief section on avoidance conditioning by saying that: 'The theory is quite unable to explain the sequence of events which we are calling introjection', but this statement may be criticized on two grounds. In the first place he has made no attempt to produce such an explanation in terms of avoidance conditioning, and to assert a universal negative of this kind is not sufficient by itself for the statement to be taken seriously. In the second place the evidence for the 'sequence of events' in question leaves much to be desired. I will concentrate on only one point which for me seems to be crucial. Argyle throughout relies on correlations between parental behaviour, values, etc., and children's behaviour, values, etc., for evidence that the former in some causal way produce the latter. As I have argued elsewhere this is one possible explanation for the existing correlations, but there are alternative explanations which are not even considered by Argyle. To take but one example, it is often quite plausible to assume that a certain type of behaviour in the child (e.g. refractoriness) may call forth a certain type of behaviour in the parent (e.g. punitiveness); such a sequence of events might be completely misrepresented by explaining an observed correlation between child's refractoriness and parent's punitiveness in terms of the latter producing the former. Even more important, in terms of the evidence cited in *Crime and Personality*, is the possibility that child's behaviour and parent's behaviour are correlated because of genetic factors. According to the genetic model the correlation between parental behaviour and child's behaviour is due to hereditary causes directly or indirectly determining both, giving rise to correlations which it would be quite erroneous to interpret as evidence of a direct causal relation between parent's behaviour and child's behaviour. It is of course customary in modern psychological writings to over-emphasize environmental factors and to disregard alternative hypotheses, but it should be emphasized that there is no scientific rationale for preferring one interpretation to another. It should be the duty of anyone dealing with correlational evidence of this kind to discuss impartially the various feasible interpretations rather than arbitrarily select the one which is more in accord with his hypothesis.

There are, I think, two criticisms which may justifiably be made of the Mowrer/Eysenck hypothesis. The first of these is that very little is in fact known about the exact details of the early upbringing of children, and the effects different methods

may have on their behaviour and their 'conscience'. It is to be hoped that the formulation of the theory will lead to more pointed research in this area. The other criticism is that the possible influence of operant conditioning (positive reinforcement of desirable activities) has been overlooked too much. There is an obvious reciprocal inhibition of 'good' and 'bad' behaviour, and the possibility cannot be gainsaid that in many cases the building up of good habits through positive reinforcement may lead to a reciprocal inhibition of bad habits, without the necessity of postulating a 'conscience' acquired through the Pavlovian type of aversion conditioning process. Given this addition to the theory, however, I feel that it can encompass satisfactorily all the known facts in this field and can lead to better experimentation than has characterized the past thirty years.

References

Argyle, M. (1964). Introjection: a form of social learning. *Brit. J. Psychol.* **55**, 391–402.
Church, R. M. (1963). The varied effects of punishment on behavior. *Psychol. Rev.* **70**, 369–402.
Eysenck, H. J. (1964). *Crime and Personality*. London: Routledge and Kegan Paul.

(*Manuscript received* 12 *January* 1965)

Brit. J. Psychol. (1965), 56, 2, *and* 3 *pp.* 309–310
Printed in Great Britain

EYSENCK'S THEORY OF THE CONSCIENCE: A REPLY

By MICHAEL ARGYLE
Institute of Experimental Psychology, University of Oxford

Eysenck (1957) put forward the theory that the conscience is based on avoidance conditioning. I put forward the view (Argyle, 1964*a*), that it is due to the learning of self-reactions, and criticized his theory. He has put forward a somewhat revised version of his theory (Eysenck, 1964), and replied (Eysenck, 1965) to my criticisms.

My criticism was that the avoidance conditioning theory is inconsistent with the finding that conscience-formation is *inversely* correlated with amount of physical punishment, and that delinquents have received *more* punishment than non-delinquents. In the recent statements of his theory Eysenck has clarified and somewhat modified his theory in a way that evades this difficulty. He now distinguishes between two kinds of punishment, and this provides an *ad hoc* explanation of these apparently discrepant findings. Aversive conditioning is a kind of learning 'which occurs very early in the child's life when a misdemeanour is *immediately* followed by a slap, withdrawal of love, or some other punishment administered within the time interval known to produce Pavlovian conditioning' (1965, p. 305). He now distinguishes this kind of punishment from heavier, more delayed punishments, to older children, which would have the reverse effect, through creating rigidity, and thus an adherence to the behaviour which was punished.

Eysenck has modified his views on a second aspect of his theory. He originally maintained that delinquents should be extraverted, since the latter condition slowly. None of the British studies with the M.P.I. have confirmed this, though there is evidence that *some* delinquents, e.g. psychopaths, are impulsive ('behavioural extraversion'). This, too, is a welcome revision, but it does limit the power of the theory—to saying that psychopaths are impulsive, and condition slowly. It does not show how psychopaths differ from other extraverts who are not delinquent.

I welcome these revisions, and I should like to restate my remaining objection. It is a sound strategy to seek a reductive explanation of complex phenomena in terms of familiar and elementary processes. However, there is a danger of overlooking the true complexity of the phenomena to be accounted for. Can the acquisition of the conscience be explained in terms of the kind of learning which governs eye-blinks? Possibly it can, but not in any simple way. The conscience consists of guilt and other complex self-reactions, not merely of restraints on particular responses. For example, children from eighteen months onwards verbalize the restraints, e.g. saying 'no' or 'don't touch', echoing parental reactions. In my paper (1964*a*) I suggested that a further explanation of this could be in terms of the learning of self-reactions from parents, either by avoidance conditioning, positive reinforcement, or verbal persuasion. In that case the two theories could be compatible, but at different levels. My objection to Eysenck's account is that it fails to show the most important aspects of the learning process here.

In fact several rather different types of social learning seem to be involved in socialization and attitude change, including identification, and various kinds of cognitive and verbal influence (Argyle, 1964*b*). It is not yet known if these can be given a reductive explanation; so far this has not been achieved. It is, however, fairly clear that they can take place without any reinforcement as this is usually understood.

Given that a number of kinds of learning *can* occur during socialization, the question is which *do* occur, and under what conditions. I have recently presented evidence (Argyle & Delin, 1965) to show that socialization processes are not universal, but only take place under certain conditions. Some require a warm relation with the mother, or only work for boys, or for ectomorphs. Eysenck has produced evidence showing that his aversive conditioning works best for introverts; I have shown that 'introjection' works best when there is a warm and dependent relationship, and with the same-sex parent.

I turn now to the general methodological points that Eysenck raises. I agree that what appear to be socialization sequences could be due to genetic resemblances. The only remedy is to study adopted children—though this is perhaps a less serious problem where, as in my study, totally different behaviour on the parts of parent and child are being related; it could not account for the findings on changes in self-perception in adults which I reviewed.

The other point is more difficult: children may elicit responses from their parents. (Parent and child constitute a dyadic system of interaction, to which each contributes, and to which no doubt the parent contributes most.) No research techniques have yet been devised which will disentangle this—apart from experiments, which are not really practicable in this field. It might be possible to study parental styles of socialization before the birth of the child, and at intervals afterwards, and likewise the development of child personality. Meanwhile what my study did was to test a prediction about the origins of self-reactions; this was confirmed for a minority of the children, and reverse causation seems in this case rather unlikely.

REFERENCES

ARGYLE, M. (1964*a*). Introjection: a form of social learning. *Brit. J. Psychol.* **56**, 391–402.
ARGYLE, M. (1964*b*). *Psychology and Social Problems*. London: Methuen.
ARGYLE, M. & DELIN (1965). Non-universal laws of socialization. *Hum. Relat.* **18**, 77–86.
EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.
EYSENCK, H. J. (1964). *Crime and Personality*. London: Routledge and Kegan Paul.
EYSENCK, H. J. (1965). A note on some criticisms of the Mowrer/Eysenck conditioning theory of conscience. *Brit. J. Psychol.* **56**, 307–9.

(*Manuscript received 31 January* 1965)

Brit. J. Psychol. (1965), **56**, 2 *and* 3, *pp.* 311–312
Printed in Great Britain

COMMENT ON W. P. BROWN'S CONCEPTION OF PERCEPTUAL DEFENCE

By E. M. COLES
Department of Psychiatry, University of British Columbia

In a monograph supplement to this *Journal*, Brown (1961) attempted to establish the thesis: 'That recognition thresholds at first rise with increases in stimulus emotionality, but reach a peak, and subsequently fall with further increases in stimulus emotionality' (Brown, 1961, p. 44). This thesis is contrary to what might have been predicted from Duffy's concept of arousal (see Duffy, 1957, p. 268; 1962, pp. 112, 123–4); and from the results of an experiment by Trehub's (1954), both of which suggest a U-shaped relation between thresholds and stimulus emotionality.

There are three possible explanations for this apparent contradiction. First, it may be argued that, although the change in recognition threshold is not in keeping with Duffy's concept of arousal, the efficiency of the defence mechanism is. Secondly, the direction of threshold change may be dependent upon the threshold as well as the degree of stimulus emotionality. Such a hypothesis has been put forward by Blum (1954). Thirdly, the experiments cited by Brown may not have used valid indicators of stimulus emotionality. It is the purpose of the present note to elaborate on this third possible explanation, and to point out that the generality of Brown's thesis is not supported by the evidence he presented.

Brown's evidence is derived from three main experiments: Bruner & Postman (1947), DeLucia & Stagner (1953) and Brown (1961); and from four confirmatory experiments: McClelland & Liberman (1949), Neel (1954), Moulton, Raphelson, Kristofferson & Atkinson (1958), and Levy (1958).

In each of the main experiments the associative reaction-time was used as the criterion of stimulus emotionality. Nowhere in his monograph did Brown establish that this is a valid criterion. In fact, there is evidence to suggest that it is not (Laffal, 1955; Singer, 1956); and Brown (1964) has, himself, subsequently stated that the associative reaction-time is not a valid indicator of 'emotional disturbance' unless combined with other measures. Thus, the confirmatory experiments, which Brown regarded as weak supporting evidence, become crucial for the generality of his thesis.

In order to establish a curvilinear hypothesis, it is necessary to measure three distinct levels of stimulus emotionality. The abstract which Brown cited as his reference to Neel's experiment suggests that her experiments provided these three levels. However, examination of her extended report (unpublished) reveals that this was not, in fact, the case.

In the three remaining confirmatory experiments, Brown has reinterpreted the original data in order to obtain evidence in support of his thesis. While it is true that an experiment can contain evidence relevant to a hypothesis other than the one the original experimenters had in mind, since the results are capable of more than one interpretation, they cannot provide *crucial* evidence for either.

It must be concluded, therefore, that, whereas Brown has demonstrated a relation between thresholds and associative reaction-time, he has not provided an unequivocal demonstration of a relation between recognition thresholds and stimulus emotionality.

References

Blum, G. S. (1954). An experimental reunion of psychoanalytic theory with perceptual vigilance and defense. *J. abnorm. soc. Psychol.* **49**, 94–8.

Brown, W. P. (1961). Conceptions of perceptual defence. *Brit. J. Psychol., Monogr. Suppl.* no. 35.

Brown, W. P. (1964). Emotional indicators in associating to neutral and emotional words. *Bull. Brit. psychol. Soc.* **17**, no. 54, 2A.

Bruner, J. S. & Postman, L. (1947). Emotional selectivity in perception and reaction. *J. Personality*, **16**, 69–77.

DeLucia, J. J. & Stagner, R. (1953). Emotional *vs.* frequency factors in word-recognition time and association time. *J. Personality*, **22**, 299–309.

Duffy, E. (1957). The psychological significance of the concept of 'arousal' or 'activation'. *Psychol. Rev.* **64**, 265–75.

Duffy, E. (1962). *Activation and Behavior*. New York: Wiley.

Laffal, J. (1955). Response faults in word association as a function of response entropy. *J. abnorm. soc. Psychol.* **50**, 265–70.

Levy, L. H. (1958). Perceptual defense in tactual perception. *J. Personality*, **26**, 467–78.

McClelland, D. C. & Liberman, A. M. (1949). The effect of need for achievement on recognition of need-related words. *J. Personality*, **18**, 236–51.

Moulton, R. W., Raphelson, A. C., Kristofferson, A. B. & Atkinson, J. W. (1958). The achievement motive and perceptual sensitivity under two conditions of motive arousal. In *Motives in Fantasy, Action and Society: a Method of Assessment and Study*, ed. Atkinson, J. W. Princeton: van Nostrand.

Neel, A. F. (1954). Conflict, recognition time, and defensive behavior. *Amer. Psychologist*, **9**, 437.

Singer, B. R. (1956). An experimental inquiry into the concept of perceptual defence. *Brit. J. Psychol.* **47**, 298–311.

Trehub, A. (1954). A theory of sensory interaction: an experimental investigation of the relationship between autonomic activity and visual sensitivity. Unpublished doctor's dissertation, Boston University.

(*Manuscript received* 19 *February* 1965)

FURTHER COMMENTS ON PERCEPTUAL DEFENCE

By W. P. BROWN

Department of Psychology, University of Aberdeen

I fully agree with Coles about the inadequacy of the evidence underlying my curvilinear hypothesis. There are, however, three points which are worth making about the present position.

If we go through the literature on perceptual defence, striking out all the evidence based on an emotional interpretation of associative reaction-time, we are left with a conflict of evidence: emotion both raises and lowers thresholds. My curvilinear hypothesis, however weakened, still provides a possible resolution for this contra-

diction. However, until we find a satisfactorily valid measure of degree of stimulus emotionality, we are in no position to test any hypothesis that is complex enough to have a chance of proving adequate.

I am not yet willing to write off associative reaction-time completely as an emotional indicator. The trouble is not that it is not sensitive to emotional disturbance, but that it is sensitive to other variables as well. In a forthcoming article which is an extended version of the abstract referred to by Coles, (Brown, 1965) I suggest that it might be possible to construct a word-association list in which long associative reaction-time would be a valid indicator of emotion. But I admit that I know of no researcher who has done this.

Finally, we seem to be left with a demonstrated curvilinear relation between thresholds and associative reaction-time. I have argued in the monograph (Brown, 1961, pp. 58–9), and I still feel, that there is no obvious way to account for this relation. It is a tantalizing problem, and it deserves to be studied in its own right.

References

Brown, W. P. (1961). Conceptions of perceptual defence. *Brit. J. Psychol., Monogr. Suppl.* no. 35.

Brown, W. P. (1965). Emotional indicators in word association. *Brit. J. Psychol.* **56**, (in the Press).

(*Manuscript received* 8 *April* 1965)

Brit. J. Psychol. (1965), 56, 2 *and* 3, *pp.* 315–319
Printed in Great Britain

APPARATUS NOTES

A SIMPLE METHOD FOR AUTOMATIC PRESENTATION OF SEQUENCES OF AUDITORY AND VISUAL STIMULI

By D. ROWAN and W. STEVENSON

Regional Physics Department, Western Regional Hospital Board, Glasgow

In psychological research it is sometimes necessary to present sequences of stimuli with various intervals of time and for various exposure times (e.g. Chapman & McGhie, 1962).

If the optimum timing is to be discovered empirically the equipment has to be flexible and the programmes easily changed. Where several types of auditory signal are to be presented, as in the distraction experiments for which the equipment was designed, the method to be described has distinct advantages.

Equipment

The requirements of the system were:

(*a*) To operate two lamps in any required sequence for any desired times of 'on' and 'off'.

(*b*) To present a distraction tone to either left or right ear in any required sequence and for any desired time, coupled with the facility of synchronization of tones with lights.

(*c*) To record the operation of lamps, distraction tone and patient's response.

(*d*) To have provision for word distraction to be applied to the patient coupled with the facility of synchronization of words with lights.

The equipment employed consists of (1) a two-channel tape recorder, (2) a six-channel event recorder, and (3) a control unit.

The Control Unit

The control circuit, shown in Fig. 1, incorporates a reed bank—an electromechanical device which is frequency selective. The reed bank consists of a single coil and ten gold-plated reeds of different resonant frequencies within the range 250–500 c/s. Only four reeds are used. The application of the appropriate resonant frequencies f_1, f_2, f_3, f_4 to the coil energizes the reeds Y_1, Y_2, Y_3, Y_4. The reeds control relays RL_1, RL_2, RL_3, RL_4 which operate lamps L_1, L_2 and earpieces Q_1, Q_2, respectively, and simultaneously actuate the event recorder. Since the energizing currents to these relays are fluctuating, reservoir capacitors C_2, C_3, C_4, C_5 are required to ensure proper operation. When relay RL_3 or RL_4 is energized the output of the 1,000 c/s transistorized oscillator is fed to the corresponding earpiece.

Operation

Fig. 2 is a block schematic diagram of the apparatus. To make up a programme the four signals f_1, f_2, f_3, f_4, which are obtained from an audio signal generator, are recorded on the tape deck. The recording times of the individual frequencies and spaces are determined by the required operations of lamps and earpieces. To operate

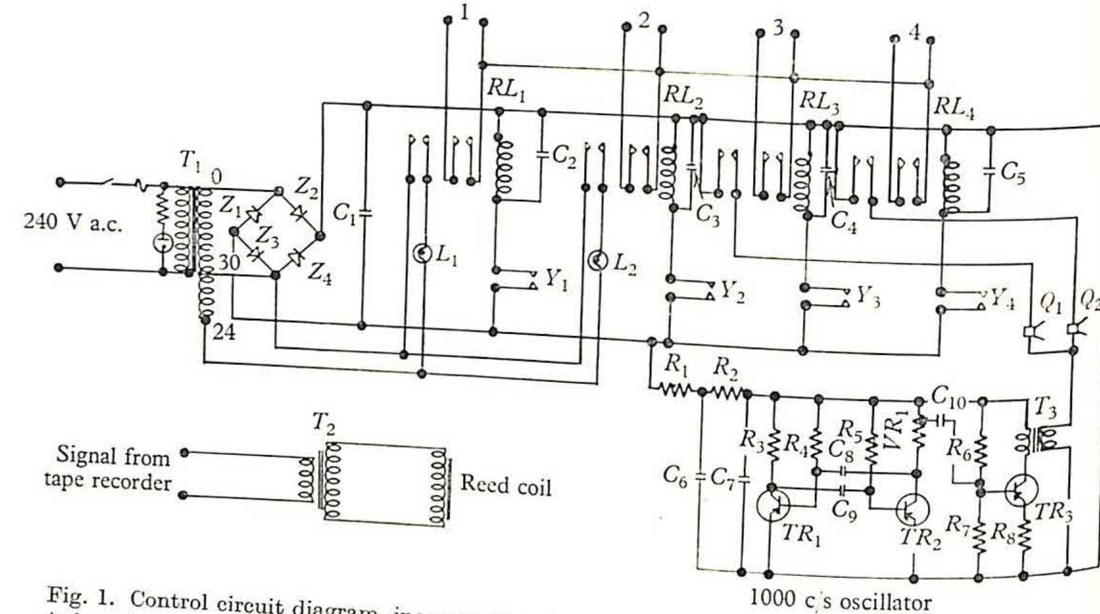


Fig. 1. Control circuit diagram, incorporating the following components: RL_1, RL_2, RL_3, RL_4 (relays, type 3000, 20 KΩ). T_1, mains transformer (Sec. 0–24–30 V, 1A); T_2, input transformer (ratio 1:40); T_3, tone output transformer (ratio 6:1). L_1, L_2 (lamps, 6·3 V, 0·3 A). TR_1, OC71; TR_2, OC71; TR_3, OC72. Reed unit, R.E.P. 100 mA). Z_1, Z_2, Z_3, Z_4 (bridge rectifiers, 50 V, 10, 160Ω (Radio and Electronic Products, Middlesex).

R_1	680 Ω	R_5	27 KΩ
R_2	680 Ω	R_6	3·3 KΩ
R_3	1 KΩ	R_7	1 KΩ
R_4	27 KΩ	R_8	68 Ω
C_1	50 MFD	C_6	50 MFD
C_2	0·5 MFD	C_7	50 MFD
C_3	0·5 MFD	C_8	0·02 MFD
C_4	0·5 MFD	C_9	0·02 MFD
C_5	0·5 MFD	C_{10}	8 MFD

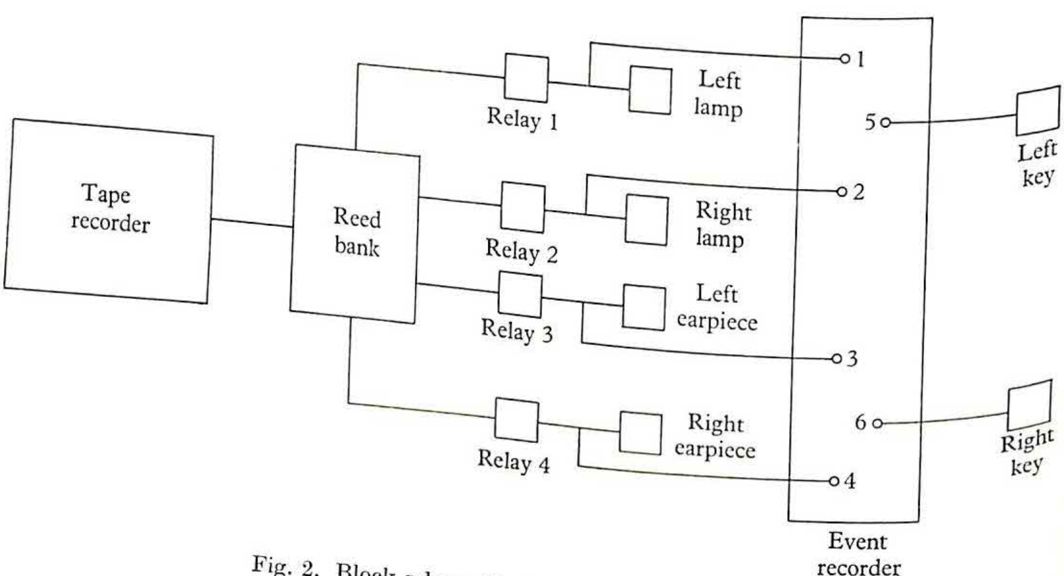


Fig. 2. Block schematic diagram of the equipment.

the lamps only, frequencies f_1 and f_2 are recorded on track 1 and the tape recorder played back on track 1. To incorporate tone distraction, frequencies f_3 and f_4 are recorded on track 2 and the tape recorder played back on track 1+2. Synchronization of tone distraction with either lamp is accomplished by recording the tone frequency on track 2 in line with the appropriate lamp frequency recorded on track 1.

For word distraction, the lamp frequencies are recorded on track 1 as before and words on track 2. The tape recorder is played back on track 1+2. In this case, the words are fed directly to both earpieces from output 2 of the tape recorder. The subject depresses a left or right morse key in response to the corresponding light stimulus, the operation of the keys being recorded on the appropriate channel of the event recorder.

A great number of programmes may be accommodated on a single reel of magnetic tape, the location of any individual programme being facilitated by the use of coloured leader tape between programmes and by reading off the tape position indicator. Unwanted programmes are easily discarded by means of the erase head of the tape deck.

This work was supported in part by a grant from the Medical Research Council. The authors wish to acknowledge the many useful discussions concerning the design of the equipment with Mr C. E. Gathercole and are grateful to Dr J. M. A. Lenihan for his constant support and encouragement.

Reference

Chapman, J. & McGhie, A. (1962). A comparative study of disordered attention in schizophrenia. *J. ment. Sci.* **108**, 487–500.

(*Manuscript received* 11 *June* 1964)

AN INEXPENSIVE DIGITAL STOP-CLOCK

By JOHN SPENCER

Department of Psychology, Bristol University

A wide variety of laboratory situations exists for which a stop-clock is an essential tool. This applies to both teaching and research situations which often only require temporary 'lash-ups' on which not much money can be expended. In these circumstances the present unit is probably adequate. It is not as accurate as a 'Chronotron' in that it only discriminates to $\frac{1}{10}$th sec, but being a counter it happily measures from 0·1 sec up to about 26 hr before it starts to repeat. This is a useful compromise where very variable cycle times have to be measured because, if a reasonable sample of cycles can be obtained, an accurate measure of the mean time/cycle can be obtained for comparisons.

It is useful if such a clock is easily portable, physically compact and mechanically robust, and it is important that such a clock should be reliable and easy to read. The clock described fulfils these desiderata reasonably well and it is easy to make. The total cost of components at current prices is under £7.

Design

The clock comprises two parts: a multivibrator and a monostable switch. Both parts use transistors and operate from a 12 V battery supply.

The *multivibrator* provides pulses of 2 msec duration at a pulse repetition frequency of 10 per sec. The circuit selected has the great advantage that temperature changes have very little effect on its operation. It is also comparatively independent of battery voltage changes. A more detailed description of the type of circuit adopted is given by Foss & Sizmur (1961).

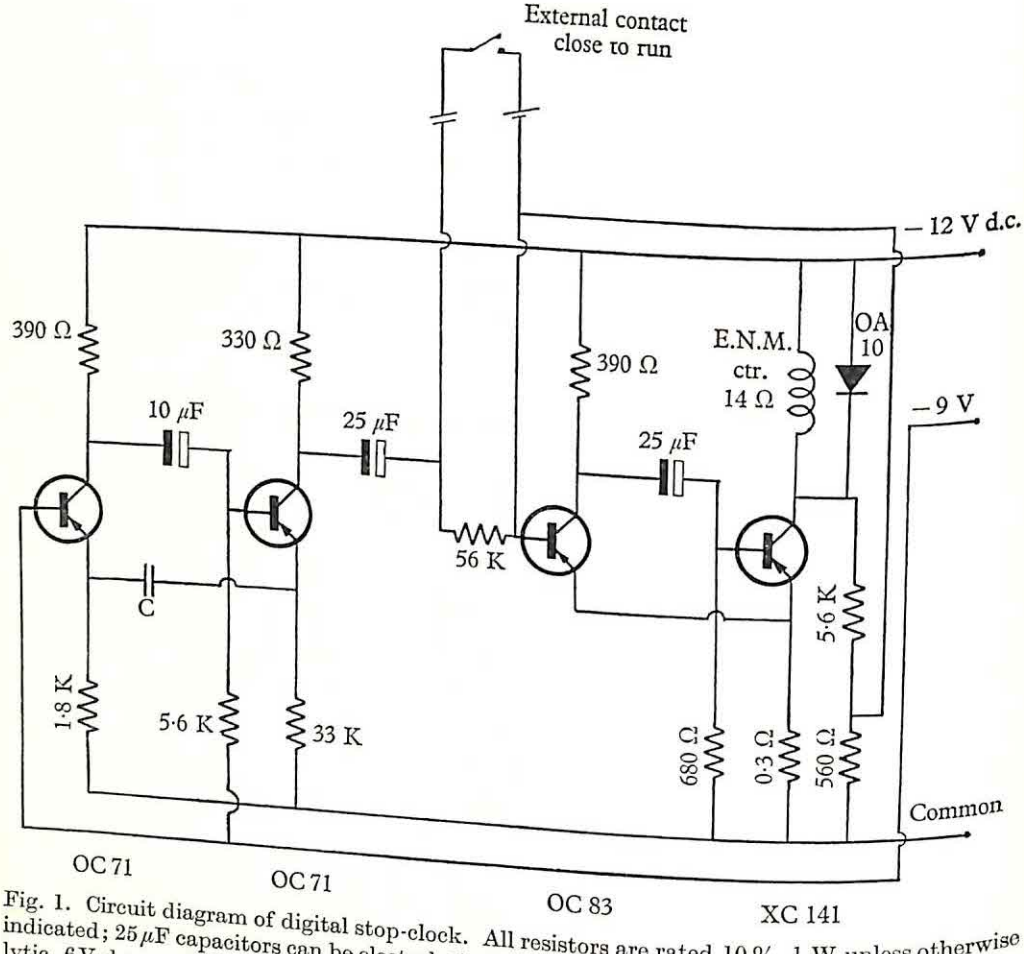


Fig. 1. Circuit diagram of digital stop-clock. All resistors are rated 10 %, ½ W unless otherwise indicated; 25 μF capacitors can be electrolytic, 15 V d.c. working; 10 μF capacitor can be electrolytic, 6 V d.c. working.

Pulses from the multivibrator are coupled to the *monostable switch* which trips for approximately 15 msec on receipt of each pulse. When the switch trips it delivers a heavy current (0·8 A) to the coil of the electro-magnetic counter. Thus, for every pulse received, the counter adds one unit to its indication. The circuit is so arranged that only during the measuring period are pulses coupled to the monostable switch. At other times the pulses are attenuated by a high resistance (56 kΩ, in Fig. 1). It is only necessary to arrange that an external contact be closed for the measuring period. This contact carries no d.c. current, and can be very small. A photo-resistor can be

used, for example, as an external contact, but the illumination changes must be sufficient to cause a large and abrupt resistance change if no amplification is provided.

The *capacitor*, C, is calculated as $19{\cdot}4\mu F$. Owing to the tolerance on electrolytic condensers, it is easier in practice to build up from several available values until the desired 10 pulses/sec are achieved. Alternatively, tantalum or paper condensers may be used to give greater stability, at greater expense with the former and greater bulk with the latter.

The *counter* is an electro-magnetic counter Type 4426–551–821 (English Numbering Machines Ltd.) with manual re-set.

The $0{\cdot}3\,\Omega$ *resistor* can be simply made by winding 1 ft. 10 in. of 36 s.w.g. enamelled copper wire onto a matchstick, and then varnishing or impregnating with polystyrene cement.

The prototype unit, excluding batteries, fits into a tin 3·25 in. diameter by 4·25 in. long, and weighs 16 oz. Quiescent current drain is approximately 33 mA. This rises during the 15 msec switch phase to about 850 mA. Accuracy of timing an accurately known duration was $\pm 1\%$ under constant environmental conditions. According to Foss & Sizmur (1961) the P.R.F. will rise by about 4% for a change in transistor temperatures of 30° C, from 20° to 50° C. Under any normal circumstances in laboratory use a change of temperature of this magnitude is extremely unlikely.

Reference

Foss, R. C. & Sizmur, M. F. (1961). Multivibrator design. *Wireless World*, May issue.

(*Manuscript received* 30 *November* 1964)

Brit. J. Psychol (1965), 56, 2 *and* 3, *pp.* 321–347
Printed in Great Britain

PUBLICATIONS REVIEWED

A Short History of British Psychology (1840–1940). By L. S. Hearnshaw. London: Methuen. 1964. Pp. vii+331. 35*s*.

Since he has recommended a group of American texts for his introductory course since the 1946 edition of Munn, the present reviewer started to read the publisher's blurb of this scholarly and clearly important work by a major British psychologist with a heart which would have sunk even more had he not remembered that its erudite author had in fact made, some years ago, some very complimentary remarks about that well-produced and, in a real sense, still comprehensive American survey. Ignoring though it did contributions from this side of the Atlantic, it was far from 'distorting' the role of British pioneers. Several generations of post-war students have been gratified and stimulated by what Prof. Hearnshaw describes as the 'modern tendency to regard the subject through American spectacles', for psychology not so very long ago was almost as dismally unenticing a science, judged by its introductory texts, as economics. Most practically and scientifically inquiring young people like to think that their intellectual adventures will cast a brighter light on the world around them and give it new colour, besides presenting them with insight into their own development and their relation to themselves, their fellows and their society. Just as Nicholson's *Elements of Economics* flung, like tablets from Sinai, Laws of Supply and Demand and definitions of high abstraction at the head of a sixth former wanting to understand the Wall Street crash, so Stout's *Manual of Psychology* had the character of the Oxbridge Don's college garden midsummer musing and was not as far away from the philosophical tradition as its author thought when he presented, boldly in his day, Hughlings Jackson's doctrine of levels of physiological processes as corresponding to organized and systematic mental activity. Professor Hearnshaw is quite right in discerning 'a new mode of psychological thinking' emerging in Stout's psychology 'in spite of its old-fashioned dress'; and it is part of his great narrative skill that, in the many complications of this interesting but well-condensed exposition, the main stream of this 'new mode' is traced from its many early springs in the nineteenth-century writings of Carpenter and Maudsley, and followed down its main tributaries in the work of Galton and Pearse, Sherrington and Head, Lloyd Morgan and Hobhouse, Graham Wallas and McDougall, to their more recent parallel courses (giving hope of eventual confluence into a single river) in the present London and Oxbridge Schools, influenced in their very different ways by Spearman and Burt on the one hand, and Myers and Bartlett on the other. The social psychology of the development of the group behavioural process that is science, the influence of original minds and strong personalities, the practical pressures of social problems, of political ideals, of war, of the rise of new sciences and of the passing cyclones and anti-cyclones of philosophy and metaphysics, all are considered in a meticulously documented manner and each of the fifteen chapters of this survey is provided with a valuable select bibliography.

The book, as such a history must be, is full of *curiosa*. We learn that the first psychologist to use analysis of variance was R. H. Thouless in assessing Whateley Carrington's work on 'trance personalities' (in the *Proceedings of the Society for Psychical Research*); that the Reading University department was started by Wolters with an annual grant of £10 and a 'very foul attic'; that the Society for the Study of Orthopsychics (1915) had as first Presidents Percy Nunn and L. T. Hobhouse, and ran a training course for nurses which included a personal analysis!

Judgements necessarily abound in a work inescapably bound up with the need to evaluate, and many of these, but not all, are agreeable to this reviewer. It is disappointing, for instance, for a one-time member of R. G. Collingwood's lecture audiences and seminars to find his subtle and friendly attitude to the development of psychology represented only by his tart criticisms of Spearman's 'cognitive' psychology. Collingwood was fond of comparing the fruitfulness of the concept of unconscious motivation for the furtherance of our understanding of human behaviour with the barrenness of 'noegenesis' and, although he may be accused now of being perhaps a little over-optimistically Freudian, who shall say he was wrong about Spearman's laws of 'creative' thinking? Certainly it would be unfair to give the impression that he 'derided' in

any way 'the modern claim of psychology', as such, 'to be a "Science"'. He felt that the science of psychology was more related to the processes of the body in which the laws of nature ran and to our feelings than to the logic of thinking, and that the progress of the science therefore could be best pursued in the less rational aspects of man's behaviour.

In other instances, one might think the judgements less than fair to Oxford philosophers. Sir David Ross and Lord Lindsay were both powerful friends of the developing discipline, all the more influential with the reactionaries because they actually read McDougall, and Ross devoted philosophical seminars to both Freud and McDougall which one remembers as exciting an interest that led one philosopher into the academic renegadism to psychology so characteristic of many of the early psychologists in this history. Sir David Ross's services on the Board of Studies in Psychology were invaluable in preparing for the establishment of the Chair, but it is sad to find that H. H. (Perception) Price is given no mention in Prof. Hearnshaw's treatment of the academic and philosophic background of psychology between the wars in this chapter XIV. Price was one of the first to introduce a critical and informed discussion of E. B. Holt's *The Freudian Wish* and encouraged all his philosophers to take a serious if critical interest in American behaviourism as much as in Freud. Perhaps, indeed, the early criticism which psychology had to meet at Oxford led to a more cautious and a more properly comprehensive experimental orientation of the discipline. Even Ryle's *The Concept of Mind* had a beneficial and astringent effect and one which is still likely to attract an inquiring mind to the placing of our 'congeries of studies' in the academic firmament—first perhaps by philosophical 'armchairmanship', but not for long, one would expect, because there are now not only many well-produced American text-books full of empirical inquiries lying about on the tables in Blackwell's, but many interesting things actually being done in the behaviour laboratories in South Parks Road.

But such small criticisms are only worth a place in an assessment of this history because it is likely to be a standard work on its subject for many years to come. It will encourage students to attain what is necessary in a human science—a sense of historical and social process—as well as exciting wonder at scientific discovery through the individual endeavour of remarkable individuals.

GEORGE WESTBY

A History of Psychology. By E. E. ESPER. Philadelphia: H. B. Saunders. 1964. Pp. 368. 45*s*.

Professor Esper sees three reasons for studying the history of psychology. First, knowledge of what has been done helps to avoid repetitions and to take advantage of possibilities that may have been suggested long ago but have lain dormant. Secondly, knowledge of our historical roots may be relevant to charting our future course. Thirdly, it contributes to the general education of scientists.

Presumably, if the third aim is to be adequately realized, some attempt should be made to encourage scientists to think historically, to develop, in relation to their own subject at least, a form of thought other than their own, as well as a body of knowledge. An attempt should be made to recreate the intellectual climates of the past and to reconstruct imaginatively how past thinkers have tried to deal with problems that confronted them. Professor Esper's work fails in this respect. It contains a mass of erudition which is used to grind axes about the importance of physiology as 'the basic science' and of Aristotle as the paradigm of past thinkers. It is not so much a history of psychology as a polemic padded with snippets about past thinkers.

Presumably one special reason for studying the history of psychology is to try to get clearer about the differences between the various forms of inquiry that historically have been gathered together under this name—to sort out what is distinctive of physiological questions, philosophical questions, sociological questions, etc., as distinct from psychological ones. Professor Esper has chapters on these different approaches but usually obfuscates the issues. We learn, for instance, that 'the problem with which Plato sought to deal in his theory of "forms" arose from the same phenomena which gave rise to the soul-ideas of primitive peoples and to dualistic conceptions down to our own times. In modern terms, the problem arises from the fact that an individual's cumulative experience is stored in coded form in his brain.' This quotation illustrates well Prof. Esper's general failure to understand both how problems arose and the types of problem that they were, combined with a tendency to think that such problems disappear if some incanta-

tion taken from 'linguistics' or physiology is chanted over them. For all his admiration for Aristotle, he missed one of Aristotle's most important contributions to psychology—his attempt to distinguish its concept from those of other sciences such as physiology and physics.

R. S. PETERS

The Acquisition of Language. Edited by U. BELLUGI and R. BROWN. Monographs of the Society for Research in Child Development. 1964. Vol. 29, no. 1. Pp. 191.

Concrete Reasoning: A Study of Intellectual Development. By JAN SMEDSLUND. Monographs of the Society for Research in Child Development. 1964. Vol. 29, no. 2 Pp. 39.

These two monographs are important ones. The first reports the proceedings of the fourth conference sponsored by the Committee on Intellectual Processes Research of the Social Science Research Council. The ground covered by the papers includes: the child's earliest grammatical system and the subsequent system with word class markers; the indication that early grammars are systematic derivations of adult grammar and that certain features of the derivative grammar can be predicted from the mean length of utterance; the growth of vocal and verbal behaviour, and the relationship between these during the first 30 months of life in order to elucidate the verbal aspects of schizophrenia occurring later; a sound spectrographic analysis of verbalizations made in the first 3 months of life; a discussion of language development in children with a variety of central nervous system deficits and peripheral deafness; and a speculative account of the relationship between mediation processes and the acquisition of language structure, together with a few other topics.

Some of the papers provide new and interesting findings, others report studies just beginning. In the reviewer's judgement, not all the papers will be proved to be equally valuable in the long run, but they are stimulating—as are the views of the formal discussants. Moreover, this fourth conference maintained the general level that we have come to expect from the three earlier ones. The participants were drawn from many disciplines, and linguistic concepts are often used. Unfortunately for the psychologist, when he tries to work in the field of linguistics he finds that there is no one point of view, for fundamental issues are now controversial. Nevertheless the monograph will give the student of language development a clear and authoritative picture of the state of much relevant research in the U.S.A. at the time of the conference, and, as such, it cannot be neglected.

Smedslund first defines the construct of concrete reasoning, and then explains how twelve methodological rules were applied to the devising of nine test items in order to ensure that, given a certain response, one could infer with maximum certainty that concrete reasoning had occurred. Moreover, the author makes clear the criteria upon which the correctness of the answers was judged. The tasks were administered individually to 120 children attending an elementary school in Colorado whose ages ranged from 4 years 3 months to 11 years 4 months.

Results show, as far as intra-item relations are concerned, that if a child gives the correct answer to any one item, it is possible to predict that he can justify his reply with a fairly high degree of confidence. But, in the matter of inter-item relationships, the difficulty of the items ranged from 122 passes to 77 passes. These latter results led Smedslund to realize that in designing his study he had failed to control both for variations in the kind of task which the subjects had to perform, and for perception, as from one item to another.

This monograph reports a thorough piece of work, well conceived and executed in the light of Smedslund's knowledge at the outset. Yet when he discusses his experiment in retrospect, he quickly points out its weaknesses and makes valuable suggestions for future research to investigate intellectual structure when certain variables have been held constant. The author learnt much from his study: so will the reader. The cover of the monograph states, '1963 award'.

K. LOVELL

Disorders of Language. Edited by A. V. S. DE REUCK and M. O'CONNOR. London: Churchill. 1964. Pp. xii + 356. 50*s*.

Recently the present reviewer reported in this *Journal* (1963, **54**, p. 379) on Osgood and Mirons' impressive *Approaches to the Study of Aphasia*. The present volume inevitably stimulates comparison with its transatlantic predecessor, and it comes off rather badly in the confrontation. The two books had a similar origin. The later one is the record of a Ciba Foundation Symposium on Aphasia, which was the outcome of a British Association (Psychology Section) meeting on Language at Norwich in 1961. The Foundation is to be congratulated on yet another significant contribution to international and interdisciplinary exchange.

Disorders of Language is European in background, although nine of the symposiasts, including four of the contributors of papers (Howes, Jakobson, Milner *et al.*, Neff) were from the U.S.A. and Canada. Six of the papers are English (Brain, Broadbent, Goldman-Eisler, Zangwill, Stengel, Ross *et al.*), two are French (Alajouanine-Lhermitte, Hecaen-Angelergues), one German (Bay), and one Russian (Luria). In some senses comparisons with Osgood and Miron are unfair. The London symposium had much less time for deliberation—three days as against six weeks—and the book, instead of attempting a coherent 'digested' transmutation of the proceedings, as Osgood and Miron did, consists simply of a verbatim reproduction of the fourteen papers presented, along with the details of each of the discussions that developed out of them. It has, therefore, no particular structure or continuity, but if it preserves some of the improvisation of the meetings themselves, it also retains their vitality.

Most of the topics discussed in the earlier volume are at least touched upon, sometimes with a lively suggestiveness and some interesting differences of cultural aura. Some aspects of the aphasic disturbance and of speech that are of most immediate interest and concern to psychologists perhaps get a slightly better run for their money, e.g. in Zangwill on 'Intelligence in aphasia', Goldman-Eisler on 'Hesitation, information, and levels of speech-production', and Broadbent on 'Perceptual and response factors in the organization of speech'. It does not detract from the value of the last paper that the party hardly seem to have been very accurately tuned in on his particular wavelength. The phonetic transcription by Ross, Clarke, and Haddock of the jargon from two interviews with a dysphasic patient is of particular interest in itself, and because it will enable some of us to consider how such recording should be done. Here, for example, as with the cinematographic recording of developmental material in infant behaviour, 'the standard antithesis arises between having everything conceivable available on the one hand, and on the other applying some sort of theory before it is collected and so losing some of the information' (Broadbent).

There will be no surprise that the symposium did not say the last word on aphasia. In some of the papers there is a distinct odour of traditionalism in the discussion of classification, and at the end of it all the chairman (Critchley) strikes a note of pessimism about the possibility of a logical classificatory system in dealing with the phenomena of aphasia. There is some divergence of view, natural though in a way unfortunate, between the clinicians on the one hand—who, of course, have a practical job to do—and the linguistic theorists and the psychologists on the other. This leads to some impatience with some of the questions concerning the psychology and the neurophysiology of speech to which the careful study of the aphasic disorders has still to make an adequate contribution. Nevertheless, the symposium was an important occasion. The report incorporates much that is of interest to psychologists in general, and particularly to those whose interests lie in language and in speech pathology.

GEORGE SETH

Rationality. By JONATHAN BENNETT. London: Routledge & Kegan Paul. 1964. Pp. viii + 119. 14*s*.

There has been altogether too much loose talk by psychologists about the essential continuity between animal and human behaviour and, equally, too much dogmatizing on the part of philosophers about their essential disparity. In this stimulating essay Mr Bennett has done his best to make amends for this by taking the concept of 'rationality', a concept which more than

any other is supposed to express the essence of what it means to be human, and submitting it to a careful and penetrating analysis. The conclusions he comes up with are rather different from any that either psychologists or philosophers have previously suggested.

The dividing line has usually been drawn in terms of whether an organism does or does not have the use of language. Bennett argues that while language is, indeed, necessary to full rationality it cannot be considered sufficient. He develops his argument by the happy device of starting with behaviour that looks both symbolic and intelligent—the so-called language of bees as elucidated by Von Frisch—but is, in fact, neither. He then proceeds, step by step, to show what would be needed if this form of communication were to count, first, as intelligent, rule-guided, linguistic behaviour and, ultimately, as truly *rational* behaviour. Before the latter could be granted, the bees, he insists, would have to display 'the sort of behaviour which can be related directly to that which is neither present nor particular', or in other words use a language capable of expressing both dated and universal judgements.

Mr Bennett's 'apian fables' have not entirely convinced me that his analysis and interpretation of the matter is the only legitimate one. For example, I am quite sure, despite the deceptive cases of honeycombs or spider-webs, that an archaeologist would not have the least difficulty deciding, solely on the basis of the kind of artifacts they bequeathed, whether a particular race of men, or of martians, were or were not rational, even though he knew nothing about their language or the sort of judgements it could express. However, psychologists can learn much from Mr Bennett; his final chapter on 'Insight' is a salutary reminder of how much confusion still surrounds some of the concepts we use so freely.

JOHN BELOFF

Categories of Human Learning. Edited by A. W. MELTON. New York and London: Academic Press. 1964. Pp. xvi + 365. 68*s*.

Laboratory studies of human learning have generated a literature which is both vast and unsatisfying. Methodological cults, theoretical traditions, and experimental persistence have deposited a formidable crust of facts around several circumscribed learning situations. On the other hand, many of these areas of inquiry remain conspicuously inbred and isolated, not only from the larger field of psychology, but also from each other. It has even been argued that most of this work is academic, in the worst sense, with little to contribute either to fundamental problems in psychology or to the management of learning in any practical setting. Whatever the case, there is no doubt that traditional laboratory students of human learning require to put their house in order. This is precisely what the present symposium aims to do. It adopts a plan which, although limited in scope, is realistic. Seven familiar regions of inquiry are selected. Each region is handed over to two acknowledged experts: one to survey it at some length and the other to comment briefly on this survey. And throughout, each author is invited to keep an eye open for points of contact between his region and the other six. The seven regions (and their main authors) are as follows: conditioning (Grant), rote verbal learning (Underwood), probability learning (Estes), short-term memory and incidental learning (Postman), concept learning (H. Kendler), skill learning (Fitts), and problem solving (Gagné).

The overall objective is to relate these regions to each other so as to create 'new and more fundamental ways of ordering our knowledge about human learning'. How far is this objective met? Not far, as Melton admits in his overview chapter on the taxonomy of human learning. Nevertheless, some welcome progress is made. Each contributor recognizes the need for more systematic integration of these seven, and other, regions. Furthermore, a number of explicit issues get highlighted as being necessary for this integration. Briefly, these include consideration of 'higher-order' activities in the analysis of all learning, as well as more careful specification of the prior history of the learner and of what he is actually doing during the learning task.

Overall, this is a book by and for serious students of human functioning. It contains some excellent and freshly up-to-date surveys, notably those by Estes, Postman and Fitts. For anyone engaged in laboratory studies of human learning, this book is certainly a must. For anyone not so engaged, it presents a fairly authoritative and comprehensive statement of what these studies are about.

I. M. L. HUNTER

Comparative Psychology: Research in Animal Behaviour. By S. C. RATNER and M. R. DENNY. Homewood, Ill.: Dorsey Press. 1964. Pp. xi + 773. $13.25.

Despite, or perhaps because of, the large literature in the field of comparative psychology, no comprehensive text has been published since 1935 when Maier and Schneirla brought out their *Principles of Comparative Psychology*. Certain specialist books have been published in recent years such as Thorpe and Zangwill's *Current Problems in Animal Behaviour*, but these in themselves do not cover the field in a broad way. This book, however, should in some measure fill the need for an organized and extensive survey of use to students and research workers. The introductory chapter provides a well reasoned and, for the student, badly needed rationale for the study of animal behaviour. Here the aims and methods of comparative psychology are concisely stated and a thumbnail sketch of its history is provided. It is a pity, however, that the authors did not discuss the animal as a model more extensively.

Two of the authors' stated aims are to provide abundant bibliographic material and to illustrate research methods for a wide variety of animal forms and behaviour processes. These they certainly achieve. The species dealt with vary from plankton to primates, and in so doing provide a rare synthesis of psychologically and ethologically orientated investigations. Of the fifty-seven papers reproduced in full about 35% are essentially the work of ethologists. Republication of papers suggests that this is a book of readings, but this is not so. The bibliographic material has been provided in four main ways: (i) references within the text which cite documentation, (ii) references within the text marked by a single asterisk which refer to studies briefly summarized in sections with a different type-face, (iii) references marked with a double asterisk and referring to the selected studies at the end of each chapter, and (iv) references given in the bibliographies at the end of each selected study. Initially this is a little confusing, though having learned the symbols the system is easy enough to follow. An investigation by Marr (1963) cited on page 362 has been omitted from the bibliography and its source remains a mystery to the reviewer.

The framework of comparative psychology which the book provides is broadly based on evolutionary theory. The authors, however, whilst saying that the theory of evolution is necessary in comparing species, do not consider that this need provide the basis for a comparative psychology. They go on in fact to say: '...comparative psychology has a number of unfulfilled objectives, and the final one is constructing and testing a general theory' p. 6). Unfortunately they do not go on to indicate the type of theory they have in mind, but perhaps this is too strong a criticism for what is essentially a source book. The twelve chapters following the introduction cover behaviour genetics, innate behaviour in vertebrates and invertebrates, hormonal factors, motivation, conditioning, instrumental learning, learning with complex cues, and finally neural factors associated with learning. Each chapter is designed to give an overview of methods and findings with the emphasis placed on empirical rather than theoretical problems. The effects of early experience, for example, are dealt with as a function of the type of early stimulation—subnormal, supernormal and changes in the typical pattern of stimulation. This is followed by a consideration of the effects on different psychological processes and finally a brief mention of some theories about the effects of early experience.

The authors achieve their stated aims, but the most disappointing feature of the book is the lack of critical material at both the empirical and theoretical levels. It would have been of more value had the selected studies been reduced by one in each chapter and the space so saved utilized in providing a critical appraisal of the material discussed.

The book provides a very extensive and well-organized review of a large and often diffuse branch of experimental psychology and is likely to be well used in libraries for a good many years.

KEVIN CONNOLLY

Foundations of Behavioural Research. By F. N. KERLINGER. New York and London: Holt, Rinehart and Wilson, Inc. 1964. Pp. xix + 739. 84*s*.

In addition to the main title this book has a subtitle, namely, *Educational and Psychological Inquiry*, but the latter need not put sociologists off, for the book is directed at a wide audience of people interested in the 'soft-wear' of behaviour research. The book is written in very general terms and the amount of verbal content far outweighs the amount of arithmetic to be digested.

It takes a lot of material to fill 739 pages, and as the book contains thirty-six chapters it is impossible to list them in a short review. The chapters, however, are divided into eight Parts, which are more or less distinct in content. Part I aims at introducing the student to the many basic concepts and constructs of the socially minded research worker in psychology in the broad sense, and also introduces elementary notions about sampling. Part II deals with what would ordinarily be thought of as 'samples', under the heading of *Sets*, and here such common concepts as means, variances and covariances are explained. In the next part elementary notions about probability and statistical inference are introduced, and there are chapters on analysis of variance, factorial designs, hypothesis-testing, and also on non-parametric procedures. When dealing with the latter concept the author uses a lot of good common sense. Part IV, which is described as 'the structural heart of the book', deals in a very general way with the design of investigations, while in the following part the distinction between laboratory-type and field-type experiments is emphasized and the writer's personal interest in experimentation of the latter type becomes clear. Part VI deals with the classic psychometric problems of measurement, reliability and validity, which are treated in a conventional way. In part VII there is a review of current methods of data collection, ranging from the interview, through methods of scaling, to chapters on sociometry, the semantic differential, and Q methodology. The final part of the book is entitled 'Analysis and interpretation'. Unfortunately, it is not an attempt to summarize or round off the content of earlier parts, but contains further methods of examining data by arrangement in contingency tables, which the author chooses to call 'crossbreaks', or by categorization in some other way. The final chapter is on factor analysis, a topic which the author no doubt felt should be mentioned, as it is so widely used in large-scale psychological and sociological studies.

In a wide general sense, this is an interesting book, if not altogether because of its actual content, at least for the many references to actual studies in the educational and socio-psychological fields. Yet this reviewer felt basically unsatisfied after reading through it. The main trouble is that the author has attempted too much. The result is that the reader gets fragments of information about an amazing variety of topics (for example, from analysis of variance to content analysis of literary material), but never enough information about any particular topic to equip him to set about some research on his own behalf. A student might benefit by reading some particular section of the book in which he was interested, provided he was then prepared to follow up the references provided.

A. E. MAXWELL

The Conditioning Therapies; the Challenge to Psychotherapy. Edited by JOSEPH WOLPE, ANDREW SABER and L. J. REYNA. New York and London: Holt, Rinehart and Winston, Inc. 1964. Pp. viii + 192. $7.00.

Conditioning Techniques in Clinical Practice and Research. Edited by C. M. FRANKS. New York: Springer. 1964. Pp. xii + 328. $8.50.

Both these books present collections of papers and manifestly have as their intention an increase in the popularity of the behaviour or conditioning therapies and techniques. The first is a report of a conference held at the University of Virginia in 1962. The editors have made an excellent job of preparing the transcript of the proceedings for publication and even the questions and answers after the presentation of the main papers do not have that slightly embarrassing flavour which occurs when direct report is made of even notable speakers' impromptu utterances. The book is subtitled 'The challenge to psychotherapy' and two chapters are taken up with a very forthright and direct example of confrontation of the psychoanalytic position. Another of the chapters is a most welcome contribution showing how experimental methods may be employed to examine at least one of the behavioural therapeutic techniques, namely desensitization. Altogether the book is easy reading, if to some extent reiterating well-worn points of view at least does it in a reasonably spritely fashion. An annotated bibliography is provided as an introduction to future reading.

An easy means of extending reading on this topic is provided by the second book, which is a collection of reprints of journal articles and chapters from books selected to cover not only therapeutic but also diagnostic uses of conditioning techniques. The papers chosen for reproduction come from a variety of English-language but in general not-psychological sources. This may

be an advantage to psychological readers in introducing them to literature with which they may be unfamiliar and to medical readers who will appreciate the respectable origins of the papers. The possible lack of knowledge of conditioning therapy in this latter audience is catered for in an introduction by the editor which runs a reasonably straight path between under- and over-simplification.

One of the virtues of both these books is their generally eclectic standpoint; neither a Skinnerian nor a Pavlovian outlook dominates. In the opening remarks to the first set of papers in the first book, the chairman, Ian Stevenson, puts forward a hope that will be re-echoed by, one hopes, all readers: that among conditioning therapists partisanship will be avoided and that there need be no compulsion to subscribe to whole systems of therapy. If these two books are anything to go by, the movement seems to be pursuing in a satisfactory inter-connecting way its several alternative directions.

PETER H. VENABLES

Pavlovian Conditioning and American Psychiatry. Symposium no. 9. Group for the Advancement of Psychiatry. New York. 1964. Pp. 211. $1.00

This pamphlet is published by the Group for the Advancement of Psychiatry, which has an active membership of about 185 American psychiatrists who are organized in the form of a number of working committees that direct their efforts towards the study of various aspects of psychiatry, and towards the application of this knowledge to the field of mental health and human relations. The present publication reports the proceedings of a conference at which two major papers were presented. The first, by Gregory Razran, is on 'The place of the conditioned reflex in psychology and psychiatry'; the second, by W. H. Bridger, is on 'Contributions of conditioning principles to psychiatry'. In addition there is some discussion of these papers by Z. Lebensohn and Ogden Lindsley.

Razran's paper, introducing the work of Sechenov, Pavlov and Bekhterev and their historical development, discusses the influences of these writers on Watson and through him on neo-behaviourism. His stimulating argument suggests that Watson may have been more correct in his views than most neo-behaviourists, particularly with respect to his thesis that the conditioned reflex is the ultimate unit of habit, which, in turn, is the basic category of behaviouristic psychology. Watson was also correct in stressing the importance of the visceral-emotional aspects of conditioning, and it is here in particular that neo-behaviourists have, in Razran's opinion, been most culpable in not following him. Razran contrasts the meagre output in this field from American laboratories with the rich harvest from Russian universities. Bridger argues for a *rapprochement* between Freudian and Pavlovian principles, and stresses particularly the importance of the second signalling system in this connexion. His contribution, like Razran's, is lively but, as Lindsley points out, it is too much concentrated on playing with words and concepts and fails to take sufficient note of actual work going on in the field of applied Pavlovian principles to psychiatry.

Psychiatrists for whom this pamphlet is written will benefit greatly from reading it; psychologists should be familiar with the material contained in it and may feel that this offering is a little too simple for them.

H. J. EYSENCK

A Psychiatric Study of Fairy Tales, Their Meaning, Origin and Usefulness. By JULIUS E. HEUSCHER. Springfield, Ill.: Charles C. Thomas. 1963. Pp. vii + 224. $7.75.

This book is a counterblast to the natural–scientific outlook which the author feels is dominating and stultifying thinking and behaviour today.

For this purpose he takes fairy tales from many different cultures and shows that they give one way of access to the meaningfulness of life, inasmuch as they deal with problems and aspects of human existence: the development of the child (chapter III) and general problems of life ('Themes', chapter IV).

The tales he considers suitable are those which have beome generalized and have lost the personal elements such as we would find in Hans Anderson. The latter kind, however, he does discuss in chapter V in the libretto of *Turandot* and its connexion with the life of Puccini.

The main part of the book presents Freudian, Jungian and 'phenomenologic' interpretations of tales, sometimes in parallel, sometimes the one which is most suited to the material. He prefers the phenomenologic as most comprehensive and classifies Freud and Jung as still in the scientific-analytic tradition. They give insight, but their view is one-sided and leaves the world meaningless.

He makes some interesting remarks on the kind of illustration a fairy tale should have (in chapter II) and some notes on Westerners and crime stories.

For anyone interested in fairy tales who has not thought of them in this kind of way, the book might be a useful, if rather expensive, introduction. A. S. PICKFORD

Eysenck Personality Inventory. By H. J. EYSENCK and SYBIL B. G. EYSENCK. University of London Press. 1964. Manual of Instructions, 24 pp., 4*s.* 6*d.* Form A or B, 5*s.* 1*d.* per doz. Scoring key, 3*s.*

The Eysenck Personality Inventory (E.P.I.) is described as 'a development of the Maudsley Personality Inventory (M.P.I.)'. Like its predecessor, it claims to measure extraversion and neuroticism but the E.P.I. comprises two parallel forms, A and B, and it includes also a 'lie scale' 'which may be used to eliminate subjects showing 'desirability response set'. We are, however, not told what to do with the 'eliminated subjects'. As Prof. Eysenck comments in a different publication, 'when people are motivated to try and give as good an account of themselves as possible, questionnaires are almost entirely useless'. Since psychological tests are used largely for purposes of selection and guidance, one might suppose that subjects do tend to be so motivated. But the corollary is not drawn in the E.P.I. manual—which reads, in fact, rather like an advertising brochure.

Thus in the section on applications of the E.P.I. we find: 'Wherever large bodies of data are being gathered the inclusion of the E.P.I. seems indicated' and, again, 'For purposes of diagnosis and treatment, the E.P.I. should be administered routinely'. In the short manual, room has even been found to denigrate other techniques: 'This whole area has been very much neglected in motivation research, where unreliable and invalid "projective techniques" have found favour'. Yet 'an enormous amount of applied work has been done in the area of market research with the M.P.I. and the E.P.I. in an attempt to discover personality correlates of consumption of different articles and brands of articles, of readership of different papers and journals, of participation in T.V., cinema going, and many other activities. The general finding has been that in almost every case where they have been looked for, these postulated personality differences have in fact been found.'

There is a tendency to assume the magic of numbers, e.g. 'The total numbers involved in all this work are in excess of 30,000'; there is also a tendency, unsuited to a test manual, to blind with science, e.g. 'T.V. viewing shows a strong, monotonic relationship with N, and a curvilinear one with E'. (According to the statisticians, these relationships are by no means mutually exclusive).

Two main points strike the reviewer. The first is the smoothness with which the Eysencks glide from 'personality tests' to personality differences, from E and N to extraversion and neurosis (as opposed to neuroticism). They attempt to draw this sting by invoking 'phenotypic and genotypic aspects' and by actually stating that their near-zero correlations between E and N 'should not be interpreted as *proving* the independence of E and N. This would imply a reification of these two conceptual entities which would be entirely inappropriate.' Appropriate or inappropriate, however, such reification is repeatedly perpetuated in other sections of the manual.

The second point is the paucity of validation data. The section called 'Validity of the scales' consists of two short paragraphs only. In the first of these it is held that since the E.P.I. scales closely resemble the M.P.I., and that validation 'exists in profusion in relation to the M.P.I. (Knapp, 1962)', therefore 'it seems reasonable to argue that this proof would also apply to the new scales. Independent proof would, of course, be required in due course, but is not yet available'. The second paragraph contains the familiar 'heads I win, tails you lose' element. Here a non-psychometric criterion was chosen, independent judges being 'asked to nominate extraverted and introverted, or stable and unstable subjects'. No figures are given but the subjects are said to 'answer the E.P.I. in a corresponding manner. There is some evidence that where there is lack of agreement, it is the judges who are at fault, rather than the inventory answers'!

Figures are given later, in two tables listing the mean scores of various normal and abnormal groups on the E.P.I. scales. Some of these figures confirm the doubts about validity already raised by the text. For example, the mean E scores of (normal) apprentices and G.P.O. telephonists are, respectively, 29·288 and 24·783, whilst those of depressives and schizophrenics are, respectively, 23·317 and 21·865. Again on the N, score, the mean of the (341) Army subjects, $m = 20{\cdot}812$, is virtually the same as the mean of the (89) schizophrenics, $m = 20{\cdot}843$.

Apart from these and other curiosities, the manual seems to have been got out in a hurry. A test–retest correlation is given as 1·736 and the heading for the last column in Tables 4 and 5 is omitted. As to the questionnaire itself, this is open to the usual psychological objections of insisting on forced choice and of telling an intelligent subject that 'there are no right or wrong answers', when the test includes such questions as 'Have you ever been late for an appointment or work?'

A. W. HEIM

Theory and Research in Projective Techniques (emphasizing the TAT). By B. I. MURSTEIN. New York and London: Wiley. Pp. xiii + 385. 64*s*.

An Index of Rorschach Responses. Studies on the Psychological Characteristics of Medical Students. 1. By CAROLINE B. THOMAS and ELLEN S. FREED. Baltimore: Johns Hopkins Press. 1964. Pp. xliii + 741. £6.

Adolescent Personality and Behavior: MMPI patterns of normal, delinquent, drop-out and other outcomes. By S. R. HATHAWAY and E. D. MONACHESI. Minneapolis: University of Minnesota Press; London: Oxford University Press. 1964. Pp. xiii + 193. 46*s*.

Children's Drawings as Measures of Intellectual Maturity. By DALE B. HARRIS. New York: Harcourt, Brace and World; London: Harrap. 1964. Pp. xiii + 367. 70*s*.

The books listed above may conveniently be considered together since they all describe extensions or advances in the use of well-established testing techniques, mainly in the field of personality assessment.

Dr Harris's book is listed last because its concern is with psychometrics rather than with personality study; in fact, the author is at pains to point out that there is as yet little evidence that children's figure drawings yield clear-cut indications for personality diagnosis, particularly when a specific task is assigned. A chapter entitled 'Clinical and Projective Uses of Children's Drawings' gives a valuable summary of work in this field, but it is incidental to the main purpose of the book, which is to present a revision and extension of the Goodenough Draw-a-Man Test. Apart from a re-standardization, the book describes two innovations. First, three drawings are now asked for instead of one: a Man, a Woman, and 'a picture of yourself'. The 'Man' and 'Woman' scores may be averaged, in terms of standard scores, or reported separately. Secondly, a quick method of evaluating drawings by means of matching with a standard series is proposed. This is the 'Quality Scale'; reliability and correlations with the full or 'Point' scale are of the order of +0·8 or better. It is also worth noting that 'Intellectual Maturity' is preferred to 'Intelligence' in the title of the book and as a definition of what is being measured. The fact that the standard scores used are scaled to resemble IQs is perhaps misleading.

The parenthetical subtitle of Murstein's book understates its preoccupation with TAT, since other projective techniques receive only passing mention, even when theoretical issues are under discussion. Viewed as a work of reference for TAT research and practice, the book is comprehensive and very adequate. In addition to discussing the broad problems of stimulus definition, validity and the like, it collates systems of analysis, assesses variations in administration and generally, provides wide and impartial coverage.

The *Index of Rorschach Responses* is a large-format volume for which most readers will have difficulty in finding a suitable place on their shelves. How useful it could be for the practising clinical or research psychologist is a matter for conjecture. It sets out a method of analysing Rorschach responses by means of KWIC (Keyword-in-Context) Indexing, described as 'a digital

computer program which retrieves all the significant words in a phrase or sentence, arranges them alphabetically, and points them out together with several words preceding and following the featured word'. The subjects were 586 medical undergraduates at Johns Hopkins (527 being 'European' males), who were tested in connexion with a 'Study of the Precursors of Hypertension and Coronary Disease'. The analysis covers Location and Content categories of the Rorschach responses, using Beck's system, and such relevant personal data as could be readily coded. Responses, which number over 20,000, and all of which are listed, are coded in terms of 'key-words'; a response may contain several key-words, and will therefore be listed the corresponding number of times. Just how one uses the resulting mass of normative data, which includes a frequency list of keywords, is not entirely clear. The standardization population is reasonably homogeneous and 'normal', but of course provides only a very dubious standard of comparison for other groups. One would have welcomed more guidance in the use of this remarkable compilation.

The background and general approach of the MMPI book listed are rather similar to those of the work just discussed. As part of a large-scale project concerned mainly with the prediction and prevention of delinquency, more than 15,000 ninth-grade school-children were tested with the MMPI and followed up over a period of 4–6 years. The results, mainly in terms of 'coded profiles', with particular emphasis on 'high-point' and 'low-point' scales, are here presented in 133 tables. These show not only associations between varieties of test performance and a wide range of relevant variables, but also a good deal of information of general sociological value, such as relationships between delinquency and 'drop-out' (i.e. premature school-leaving) and socio-economic status, community size etc. Since the populations studied are highly representative—a virtually complete city sample and a 50% 'state-wide' sample are treated separately—the results would appear to have high generalization value. The authors apologize for their use of MMPI 'jargon', but this is not excessive; given a minimal acquaintance with the technical language, one finds the discussion readable and stimulating.

B. SEMEONOFF

Contributions to Mathematical Psychology. Edited by N. FREDERIKSEN and H. GULLIKSEN. New York: Holt, Rinehart and Winston, Inc. 1964. Pp. x+189. Price not stated.

In 1962 the Educational Testing Service, which, from its foundation onward, owed so much to Leon Thurstone and those who were trained by him, decided to name its latest building the 'Thurstone Hall'. The ceremony of dedication included a conference to which half a dozen eminent psychologists were invited. All were either students of Thurstone or students of his students. Their papers, each of which deals with a topic to which Thurstone had already made a 'pioneering contribution', are now printed in the present volume.

British students frequently ask 'what is this new American science which has been christened "mathematical psychology"?' The answer commonly given is 'that branch of psychology which is so mathematical that no psychologist can understand it'. As interpreted here, it seems to cover, not so much the relatively abstruse techniques recently discussed under that name, but rather what an older generation called 'quantitative psychology', Brown and Thomson renamed 'statistical psychology' ('because all the methods are more or less statistical'), and Spearman and Thurstone called 'psychometrics'.

The first paper, by Clyde Coombs, on 'Symmetries and Dualities among Measurement Data Matrices', extends and illustrates the method of classifying such data that he has already put forward. The contributions of Gulliksen and Abelson are concerned with the study of attitudes. Guilford gives a succinct account of his views on 'creativity' and relates it to his well-known 'structure-of-intellect model'. Horst discusses 'matrix factoring and test theory', and Ledyard Tucker describes an instructive extension of factor analysis to 'three-dimensional matrices'. All the papers, except the first, involve applications or amplifications of Thurstone's factorial theories; and each deserves the close attention of every specialist who is engaged in applying such procedures to problems of individual psychology.

For the general psychologist by far the most interesting chapter is Dr Dorothy Adkins's biographical sketch of Thurstone's life and writings—a reprint of the brochure distributed on

the occasion of the conference. To it is appended a complete bibliography of his publications. Thurstone's original training was as a mechanical and electrical engineer; and Dr Adkins has several illuminating anecdotes to record about his early career. One of his first inventions was 'a motion-picture camera and projector'. This greatly impressed Thomas Edison, who gave him his first appointment as laboratory assistant. As Dr Adkins observes, 'this opportunity for close association with a genius doubtless played some part in his interest in "creativity" and his tests for "ideational fluency" and "spatial abilities"'. During World War I Walter Bingham invited him to accept a post in the Division of Applied Psychology at the Carnegie Institute of Technology. Bingham was a frequent visitor to this country, and became particularly interested in British researches on intelligence testing and factorial techniques. The experience thus gained while working under Bingham is reflected in two papers by Thurstone which were published in this *Journal*, on intelligence and on vocational guidance. These were delivered during the International Congress on Psychology at Oxford; and it was there that I first met him. They were followed by a long series of researches on psychological methods and techniques for scaling; and it was not until 1931 that he published his first paper on factor analysis—a topic which up till then had been looked on with disfavour in the U.S.A. As Dr Adkins observes, it was chiefly his work on factor analysis that received most attention in his lifetime (and, one might say, in this particular volume); but, she adds, 'what gave Thurstone himself the greatest satisfaction was his success in breathing new life into psychophysics—an evaluation in which later historians may concur'.

Perhaps I may be permitted to append a brief tribute of my own and record a further incident which may throw some light on his eager but critical approach to new ideas. In 1924, after the meeting of the British Association in Toronto, a small party of scientists arranged a trip to Niagara. Thurstone, Flugel and I were, if I remember rightly, the only psychologists. Though four years younger than Flugel and myself, Thurstone had already published far more papers than both of us put together. While still a high-school pupil of 18, he contributed a long letter to the *Scientific American* on 'How to Save Niagara'; and we naturally hoped to hear an up-to-date account of his scheme. He, however, was far more anxious to learn how we could possibly defend 'the British notion that all intellectual processes depend on a single general factor'. Flugel (in those days Spearman's chief assistant) gave a brief but brilliant exposition of Spearman's theory, and pointed out that, as propounded by Galton, 'the British notion' included *both* an innate 'general ability' *and* a number of 'special abilities' of the type that Thurstone himself accepted. I tried to explain how Pearson's original method of principal axes (a method that had long been used in celestial mechanics) envisaged 'a multiplicity of factors'—a view that seemed confirmed by my own tentative studies of school children, and by those of Sheppard (author of the well-known tables of the normal distribution and of 'Sheppard's correction'). Sheppard was one of His Majesty's Inspectors of Schools, and had just published a book which showed how the mathematical problems of factor analysis could most readily be solved by using matrix algebra instead of Pearson's cumbersome determinantal methods. Thurstone, however, thought it a pity that Spearman and Brown (Spearman's most vigorous critic) had not stuck to the problem of applying psychophysical principles to mental measurement, since in their earlier writings they had both made such valuable contributions to these well worn techniques. And, as we all three stood watching the Falls, he went on to criticize the methods I myself had used in *Mental and Scholastic Tests*—a series of criticisms to which I myself was deeply indebted in my later investigations.

A year or two later, Truman Lee Kelly, who had studied under Karl Pearson, published his *Crossroads in the Mind of Man*, in which he 'prophesied that forthcoming data, as well as much of the earlier, will show the need for more than Spearman's assumption of a single general factor' and outlined an alternative procedure of his own. In passing he referred to Thurstone's paper on intelligence, and that no doubt led Thurstone to look once again at the problem of factorization. Dr Adkins relates how in 1932 he 'approached the Carnegie Corporation with a request for funds to try out ideas on multiple factor analysis, making it clear that he could provide no assurance that the gamble he proposed would be successful'. The 'gamble' succeeded beyond all expectation. During the next ten years or so no book was as frequently quoted or used as his *Vectors of Mind*. I was delighted to see my formula of 1917 appearing as 'the fundamental factor equation'—the more so as the deduction was, I believe, entirely independent. But the outstanding merit of the book was the systematic development of the whole technique, and the lucid and logical exposition which accompanied the algebraic proofs.

Its lasting influence is still discernible in the pages of this volume; and every psychologist will heartily agree that the tributes here paid to 'one of the outstanding figures in the history of psychology' are fully deserved.

CYRIL BURT

Progress in Experimental Personality Research. Vol. I. Edited by BRENDAN A. MAHER. New York and London: Academic Press. 1964. Pp. x + 368. 36*s*.

The seven contributors to this book review work in their special area of interest and describe their own investigations.

'Input dysfunction in schizophrenia', by P. H. Venables, contains a cohesive review of the literature which refreshingly pays equally respectful attention to papers presenting careful observations (such as those of McGhie and Chapman), theories (such as that of Fish) and experiments (such as those of Shakow and his colleagues). His well-supported arguments leave little doubt that future studies must treat separately chronics and acutes, paranoids and, at least, non-paranoids, level of on-going physiological activity and reactivity to stimuli. Venables' statement: 'That over-inclusion occurs in schizophrenics with good prognosis...fits the ideas put forward earlier...' should substitute 'if' for 'that', since Payne's suggestion that over-inclusion may be a relatively good long-term prognostic indicator is based on the fact that chronics do not over-include and is fallacious because it draws a longitudinal conclusion from cross-sectional data. Acutes may lose their over-inclusiveness when they become chronic.

'A theory of verbal behaviour in schizophrenia', by L. J. Chapman, J. P. Chapman and G. A. Miller, describes L. J. Chapman's abandonment of 'over-inclusion' in favour of a more complex theory which it is hoped will account for both the over-inclusion and over-exclusion which has been found experimentally by manipulating experimental instructions. The reported findings support the view that the overt responses of schizophrenics, more than those of normal subjects, are mediated by the strongest mediating responses, whether they be aroused by the stimulus words or by contextual cues.

A review of work on 'The effect of the experimenter on the results of psychological research', by R. Rosenthal, leads him to conclude that: 'Perhaps the most compelling and the most general conclusion to be drawn is that human beings can engage in highly effective and influential unprogrammed and unintended communication with one another. The subtlety of this communication is such that casual observation of human dyads is unlikely to reveal the nature of this communication process.' He discusses how to get beyond casual observation. The implications should be disturbing, particularly for physicomimetic psychologists.

'Personality and conditioning', by S. H. Lovibond, sometimes leaves this well-trodden path to explore the less widely known recent Soviet research. The unconscious conflict theory of neurosis is naturally given less attention than the behaviour therapist's, 'neurosis is nothing but symptoms', and what Lovibond regards as the Pavlovian or Hebbian view that neurotic behaviour has to be distinguished from a hypothesized disturbance of nervous activity which constitutes the neurosis. Unfortunately, Hebb's six criteria for human neurosis apply equally well to psychosis. This apart and letting ANA = abnormal nervous activity and NB = neurotic behaviour, how can it be established that ANA constitutes the neurosis? If ANA and NB have a Spinozian synchrony, no grounds exist for any priority. If ANA can be produced which eventuates in NB, we cannot identify the N in NB without recourse to criteria other than that B was preceded by ANA. The fact that many say 'It's my nerves, doctor' should not be taken too literally.

In 'Repression-sensitization as a dimension of personality', Donn Byrne suggests that the majority of evidence supports a linear rather than a curvilinear relationship with maladjustment, but that both relata are frequently measured by self-report techniques in each of which denial may be operating. It is clearly time to concentrate more effort on measurement at different levels of awareness. Some studies indicate that sensitizers respond with higher anxiety than repressers on a self-report measure; whereas, on a more concealed measure, repressers show greater disturbance. This is much more meaningful clinically than the linear relationship.

'Interpersonal perception', by V. B. Cline, describes attempts to overcome the methodological weaknesses pointed out by Cronbach and others. Cline concludes that judging accuracy probably involves several orthogonal factors, but that one can still talk of overall ability to judge

others. Although he believes that a fully satisfactory measure of judging accuracy is still lacking, he is able to point to ways in which advances might be made.

In 'Stimulus functions in projective techniques', D. T. Kenny pleads for more attention to stimulus properties. He puts forward several postulates which seem to be susceptible to experimental confirmation or rejection. The interesting possibility emerges that attributes ascribed to figures similar to the subject are more likely to indicate conscious states; whereas, those ascribed to dissimilar figures are more likely to indicate unconscious states.

Experimentalists not specifically concerned with personality research should find the book much less foreign than they might fear; specialists in one of the areas treated may discover sufficient interlocking of topics to encourage them to have this useful, well-produced, but expensive book on their shelves.

G. A. FOULDS

Experiments in Behaviour Therapy. Edited by H. J. EYSENCK. London: Pergamon Press. 1964. Pp. x+550. £5.

This book is a successor to *Behaviour Therapy and Neuroses* (Eysenck, 1960), and is an attempt to increase knowledge appertaining to the theory and application of behaviour therapy. Like the previous book, it is a symposium, and most of the contributions have been published elsewhere, though some appear here for the first time. Approximately half of the articles are reports on the application of various methods of treatment to actual cases. The remainder consists of purely theoretical papers, reports on control trials, and of attempts to demonstrate experimentally what kinds of behaviour can be brought under the control of environmental manipulanda.

The book consists of four parts. The first deals with the method of reciprocal inhibition. It comprises interesting applications and adaptations of the method of single cases and groups of cases with a variety of psychiatric problems. On the research side three contributions deserve mention: Lazarus's comparative study of group therapy of phobic disorders by systematic desensitization with a more traditional form of interpretive group psychotherapy; Long and Lazovik's experimental investigation of systematic desensitization; and Walton and Mather's attempt to test certain inferences derived from learning theory in relation to treatment of chronic and acute obsessive-compulsive states. This section also contains three articles on aversive therapy. The paper by Blakemore is of particular interest. He draws attention to a variety of theoretical and practical problems involved in this type of treatment and offers fruitful lines for further research. Eysenck's two lucid theoretical papers, discussing spontaneous remission, transference, extinction and relapse in neurosis, complete the section.

The second part deals with operant conditioning. Ferster attempts in the opening paper to analyse abnormal behaviour in terms of reinforcements immediately available in the environment and suggests various ways by which behavioural repertoires can be generated, differentiated and eliminated. The remainder of the contributions make it clear to the reader that the method has so far been used mainly to exert control over specific behavioural anomalies in experimental settings (particularly of psychotic patients), rather than on attempts to treat disorders as such. Two papers by Ayllon demonstrate the usefulness of the methods for the purpose of modifying behaviour problems of chronic psychotic patients in hospital settings. It would appear that the generalization and permanence of behaviour changes brought about by the methods are not very impressive.

Other methods of treatment are presented in the third part of the book. Here the reader will find the application of massed practice to treatment of tics; Ellis's rational psychotherapy; and hypnosis. The latter technique is included on the assumption that it may be regarded as a phenomenon dependent on verbal conditioning. Edward's critical review of the hypnotic treatment of asthma and Lazarus's clinical applications of autohypnosis fall into this subsection. The reader will also find a comparison of the efficacy of rational psychotherapy, psychoanalytically oriented psychotherapy and orthodox psychoanalysis, and a control trial of group psychotherapy and an activity programme given to chronic schizophrenics. In addition, there is an interesting article which attempts to conceptualize asthma as a learning response, and a review of empirical and theoretical approaches to speech in normals and stammerers.

The last part deals with modification of behaviour in children. It will become apparent to the reader that behaviour therapy is rarely practised in child treatment. The section starts with Rachman's article on the application of behaviour therapy to clinical problems and therapeutic

possibilities. Apart from Lazarus's article, which illustrates the application of various methods, the remaining contributions concern themselves with the use of the procedures of operant conditioning. As with adults, the emphasis is on the analysis and control of specific forms of behaviour. Of particular interest, as regards clinical implications, are two articles by Lovaas, who demonstrates an interaction between verbal and non-verbal operant behaviour. Apart from these, the adaptation of systematic desensitization (emotive imagery), developed by Lazarus, merits attention.

This book is an improvement on the previous one inasmuch as the emphasis has shifted from illustrations of the methods used to the theoretical and practical problems entailed in the application of behaviour therapy. As one would expect, the original enthusiasm due to encouraging results has been subdued a little and a more critical tone is now more apparent. It is pleasing to note that there is a greater awareness of large gaps in knowledge and of the necessity for rigorous application of scientific principles to a variety of problems, particularly with regard to the improvement of the existing methods and the widening of their scope. This trend is particularly evident in the sections by Eysenck. Also, the greater proportion of theoretical discussions and research projects—rudimentary as they are—is a healthy sign. Some omissions are noticeable, e.g. recent work by Lindsley and Skinner.

The book brings together a great many reports, which is of tremendous value in itself. Practising behaviour therapists and research-minded workers in this field in particular will find the book of immense help. It not only provides valuable guides to application of behavioural methods in therapy, but opens up numerous avenues for research and further developments. Despite its theoretical orientation and approach to psychiatric problems, those with different views will undoubtedly profit from reading the book. But since this collection of papers was made, many new articles have appeared and because of this, the book cannot be accepted as completely adequate coverage of current knowledge. The price of the book is extremely high and there are numerous typographical errors. VICTOR MEYER

Personality Change. Edited by P. WORCHEL and D. BYRNE. New York and London: Wiley. 1964. Pp. xi + 616. £5 5*s*.

Wherever personality is seriously investigated, argument hinges first on definition. But if, as is usually the case, the definition which emerges represents personality as the enduring aspect of our being that survives all pressures and determines individual variability of reaction to identical stimuli, then there remains the persistent problem which faces all teachers of personality psychology, namely, how to reconcile this concept with the evidence of mutability at least in the superficial manifestations of the personality.

As a stimulus to graduate discussion, the editors, P. Worchel and D. Byrne, have collected papers from contributors who individually presented their material in seminars conducted at the University of Texas during 1961 and 1962. It is somewhat surprising that personality change has not been more fully documented previously and we must welcome this volume for its attempt to deal with this most intractable of the problems of personality measurement. Even so, the group of invited contributors does not entirely adequately represent all personality theories. Sheldon, for example, is ignored and the factor theories of Cattell and Eysenck receive scant attention. Furthermore, though it is consistently implied that there is an inherited substrate to personality, this aspect receives inadequate treatment. We also miss a discussion of personality change following operation or illness, e.g. changes due to brain damage or to cerebral and other operations.

We are, however, offered discussion and review of work reflecting planned and unplanned personality change due to most of the other everyday and some of the more *outré* situations and occurrences in the human life span from early personality development (L. Luborsky and J. Schimek present the orthodox psychoanalytic view with a respectable veneer of science) through to old age (ably reviewed by R. G. Kuhlen). J. L. and Marian Yarrow review personality change and interaction in its family context and find continuity compatible with change, though there is evidence in W. Madsen's study of cross-cultural transfer that this combination produces conflict. Among other notable chapters are those on change in attitudes (R. Stagner)—particularly ethnocentrism resulting in international tension, a social scientific topic which may be literally a matter of life and death—and on moral values (L. Festinger and J. L. Freedman) in which

cognitive dissonance theorists continue to achieve plausibility of argument and ingenuity of experimentation despite the recent and devastating critique by Chapanis and Chapanis (1964). Changes due to voluntary and enforced isolation are compared (E. A. Haggard) and R. R. Holt deals with indoctrination and brain washing. J. Zubin and M. M. Katz provide a stimulating chapter on drug-induced changes in personality. They distinguish physiological and conceptual components and forsee future progress in personality investigation through psychochemistry rather than psychophysics.

But it is in relation to the study of psychotherapy—what Miller on p. 150 calls 'remedial personality education'—that this book is memorable. Following a brilliant analysis of problems of measurement and alteration in personality study (including the influence of biases attributable to the personality of the therapist or personality theorist) (by D. Byrne), there are chapters which deal with various aspects of planned personality change or psychotherapy. T. R. Sarbin (on Role Theory) and E. J. Murray (on the Sociotropic Learning approach) are noteworthy but the wide-ranging and clearly illustrated chapter by N. E. Miller is outstanding. He uses recent experimental work to extend the application of the well-known Miller–Dollard theory to new psychotherapeutic situations and adds a note on psychological homeostasis and resistance to change. One wonders why there has been so little serious application of learning theory to traditional psychotherapy in this country. Despite doubts as to its effectiveness, psychotherapy is here to stay and the authors gather together and review the sound experimental work which has been carried out in this field as well as indicating the scope for further work and the difficulties, methodological and theoretical, to be overcome.

If personality study still requires justification in the scientific world, this book (and in particular the chapter by I. E. Farber), will contribute towards its establishment on a sound footing. There is much to interest the anthropologist, sociologist, psychometrist, and, as the editors hope, the literate layman. In addition, by skilful insertion of introductory and linking sections, the editors succeed in providing a collection of cohesive, if not unified, views on a major and neglected topic and certainly present essential reading for anyone engaged in psychotherapy or research into its effects.

ANNE BROADHURST

CHAPANIS, NATALIA P. & CHAPANIS, A. (1964). Cognitive dissonance: five years later. *Psychol. Bull.* **61**, 1–22.

✓ *Crime and Personality.* By H. J. EYSENCK. London: Routledge and Kegan Paul. 1964. Pp. 204. 25*s*.

Perhaps one of the chief attractions of this work by Eysenck is its herculean attempt to fit various aspects of deviant behaviour into the totality of a personality theory, at a time when forensic psychological literature is either mainly dominated by Freud's all embracing personality theory or by endless fragmented pieces of research which study isolated criminological phenomena without having the courage to link the findings to the totality of a personality theory. Eysenck has never lacked courage and his zeal has left British psychology under no illusion as to how he thinks it ought to move from piecemeal research towards a behaviourist-based personality theory.

It is within this context that his preoccupation with the two-factor theory of extraversion/introversion and neuroticism in relation to delinquency must be seen. Superficial reading of his book creates the erroneous belief that Eysenck claims that all delinquents are automatically extreme extraverted neurotics. However, he points out that quick intraverted conditioners without much neurosis may at times equally well prove delinquent, especially if exposed to a strong criminogenic environment.

A second attraction for many readers may lie in the fact that Eysenck refuses to evade the issue of heredity and criminality in an epoch where the pendulum has swung from a Lombrosian genetic approach to the prevalent belief, shared by most psychoanalysts and behaviourists, that it is more fruitful to concentrate on environmental rather than genetic aspects of the theories of causation. Eysenck, in an attempt to redress the balance, skilfully explores the avenues in both directions. On the genetic side, he quotes, among others, Shield's research on monozygotic twins, reared both together and apart, which tends to suggest that some separated twins, despite different environmental influences, appear to be genetically prone to criminality by

comparison with non-separated twins and ordinary siblings. Also he quotes Soloman's conditioning experiments where different puppies (both breeds and individuals), in response to not being allowed to eat a certain type of meat, condition more readily than others when punished. Some puppies violated the newly acquired taboo within 6 min. Others did not do so even after 16 days and over, having evidently developed a discriminating 'puppy-conscience' which, in some, seemed a mixture of deep-seated genuinely felt guilt and a more superficial resistance to temptation. This behaviour is equated with that of humans.

Eysenck and his co-workers again stress that extraverts (and, by implication, delinquents, especially if also neurotic) build up inhibition quickly, show high degrees of inhibition, and dissipate inhibitions slowly; whereas introverts build up inhibitions more slowly and dissipate it more quickly. Correspondingly, extraverts develop excitation in the nervous system more slowly and weakly than introverts who are thought to do so more quickly and strongly. He quotes Spielman, the Franks' and others who found that extraverted behaviour was often prevalent among persons who were either brain damaged or delinquent, and links this with the function of the central nervous system. This is in contrast with the neurotic induced behaviour which is tentatively linked with the autonomic nervous system (sympathetic and parasympathetic) and the as yet not well-understood brain function of the ascending reticular formation.

The third attraction for many readers may lie in the fact that Eysenck, together with only a handful of others, has attempted to bring a fresh approach to the potentially promising delinquency treatment programmes through the specialized application of reversed and non-reversed forms of reciprocal inhibition therapy. Eysenck argues that a sympathetic response to a given stimulus can be eliminated by conditioning a parasympathetic response to the same stimulus, with the parasympathetic response being antagonistic to the sympathetic one which will inhibit it reciprocally (via the distance gradient technique) and thus cancel it out. Wolpe and others are cited as having cured some delinquents, fetishists and others by using negative stimuli like apomorphine or electric shock, and by using the reverse of reciprocal inhibition technique.

Eysenck asks his readers to take a hard look at delinquency therapy, on the assumption that most delinquents are slow conditioners, often with a high degree of partly inherited and partly acquired neuroticism (via anxiety—following Mowrer's work). He calls for more research to substantiate the theory that many delinquents can probably be cured either through newly developed drugs designed to reduce neuroticism, or by drugs (especially stimulants, like the amphetamines) to counterbalance inherited components of excessive extraversion among slow conditioners. He makes a plea for more research into the efficiency of treatment based on reciprocal inhibition techniques as a form of behaviour and deconditioning therapy.

It is no use arguing that Eysenck may be wrong about several vital points. There are many who disagree with him that most delinquents are either neurotic or permanently slow conditioners, but contend that they may vary in responses to different stimuli as would a cross-section of non-delinquents. Others will question the success of behaviour therapy with complex behaviour syndromes, especially at the ideational level, and point out that too much has been made even of the more carefully controlled existing twin studies. Finally, many emphasize their view that psychotherapy is only a form of behaviour therapy, but at a subtle trial and error level involving a verbal secondary signal system.

Criminological research and treatment have so far lagged sadly behind the success of the normative sciences. Eysenck must surely be given credit for having consistently presented criminologists and forensic psychologists with a new framework, alongside the existing ones, within which to experiment in accordance with the belief that punishment must not be made to fit the crime, but that varying forms of treatment for different types of offenders must be evolved in a rational penal policy.

R. G. ANDRY

Pavlov's Typology. Compiled, edited and translated by J. A. GRAY. Oxford: Pergamon Press. 1964. Pp. xv + 480. 84*s*.

On page 289, Dr Gray modestly suggests that Pavlovian theory is fairly familiar to western psychologists. He adds that it is not usually used by them as a working tool. My impression is, however, that although some experts know Anrep and Konorski, and are familiar with Gantt's translations, the refinements of Pavlovian theory are a closed book to most. The reason for this, if I am right, is that practically nothing of Pavlov's or Teplov's work outside Anrep and Gantt

and one or two essays by Pavlov and Teplov have been available in English. For the majority, therefore, the complexities of the associations between internal and external inhibition have scarcely been sampled. This is even more true of the irradiation of excitation and inhibition and especially so for the theory of types of nervous system. Teplov's work has been even more of an unknown quantity because it was not available in English until about 1961 and then in an incomplete form. In any case his present essay is such a careful historical account of the problems of the association between strength, mobility and equilibrium of nervous reactivity, that the ordinary casual reader might well be excused for giving up the ghost.

For these reasons, if for no other, Gray's book is valuable in providing good basic material. However, it does a great deal more than provide excellently translated and edited texts of Teplov and his colleagues. It amounts to a comparative study of the usage of terms from Pavlov's earliest thoughts on the subject of nervous types up to the most recent views of Rozhdestvenskaya, Nebylitsyn and Yermolayeva-Tomina. In addition to acting as a guide through this difficult territory, the author adds two essays of his own which attempt to relate the Teplovian system to a system more acceptable to English and American psychologists. In attempting all this Dr Gray has produced a work of careful scholarship which is likely to become standard reading for anyone interested in acquiring competence in the type of study discussed.

The book is roughly divided between about 270 pages of translated text and 210 pages of commentary and exposition. The first essay by Teplov is already so much fuller and more detailed than previously available essays that it goes a long way to clearing up the difficulties readers might have had in deciding whether mobility of the nervous processes was similar to equilibrium or not, or whether an excitable animal could be behaviourally placid. In a very detailed and almost too careful manner Teplov traces the history of Pavlov's experiments and of his thinking about the meaning of 'weak' and 'strong' nervous systems and how the notion of equilibrium was introduced in about 1925 or 1926. Generally speaking typology until this time had been based on observation of behaviour in the stand and elsewhere but now Pavlov and his colleagues started to use measures such as the speed of the formation of conditioned reflexes and their relative stability. Even at this time apparently Pavlov used the limit of stimulation as a measure of strength. By about 1933 the final, third factor had been added to the system and mobility was introduced. It is the accident of the historical development of these factors one after another, one sometimes confused with another, which makes some parts of Anrep so difficult. We can get a great deal of enlightenment from working through this essay of Teplov's even though Teplov's style is scarcely that of a popular teacher. Gray's footnotes add a good deal to an already interesting text especially, for example, in a note which occurs later in Part 2 where Gray points out (p. 274) that 'strength of excitatory process' may have a different meaning from 'strength of nervous system', and in fact may mean just the opposite. Many other interesting observations occur in this and in the author's following essay which make sense of a rather confused system of variables. The reader who wants to understand what the more recent work in this field is all about might do well to start with a very good summary on pp. 274–282 and then read both backwards and forwards.

I hope that this book will be read carefully and will lead to some association between workers in this field in Moscow and similar workers both here and in the U.S.A. Dr Gray would be in a very favourable position to begin such collaborative work.

NEIL O'CONNOR

Imprinting and Early Learning. By W. Sluckin. London: Methuen. 1965. Pp. 147. 25s.

Imprinting, the change which occurs when a young chick or duckling forms an attachment to its parent or parent-substitute, was said by Lorenz to be fundamentally different from other forms of learning. His radical views gave rise to a still-growing volume of research, and though Lorenz himself now regards imprinting as 'one form of conditioning', albeit with its own peculiar characteristics, other workers have been ready to take up the defence of his earlier distinction. The controversy has been a fertile one, linking imprinting to other aspects of behaviour study: recently considerable clarification has been achieved, and a comprehensive review is clearly timely.

Dr Sluckin, who with his colleague Dr Salzen has played a leading part in this field, is well qualified to make such a review. In a survey which is both comprehensive and critical he rightly

emphasizes that to attach one or the other label to imprinting matters less than to specify precisely how imprinting compares with other forms of learning.

The recent clarification has come largely from a deeper understanding of two issues. One of the characteristics of imprinting stressed earlier by Lorenz was that its occurrence is limited to a sensitive period. The end of this sensitive period is associated with an increase in fear or anxiety responses, and the appearance of these is a consequence of experience. The second issue concerns the circumstances in which imprinting occurs. It is clearly independent of any of the conventional reinforcements, and Hess's suggestion that the strength of imprinting is related to the muscular effort expended never had any clear evidence. But, only with the appreciation of the importance of merely perceiving and learning the characteristics of the object did it become possible to understand both the sensitive period and the appearance of fear. If imprinting involves a learning of the characteristics of the environment, including the object followed, and strange objects are subsequently avoided, relationships can be seen between imprinting, 'perceptual' learning, and the development of fear and anxiety: the study of imprinting thus has repercussions on much broader issues than the following reponse of nidifugous birds. Dr Sluckin's review is to be recommended to all interested in learning and the development of behaviour.

R. A. HINDE

The Psychoanalytical Treatment of Children. By ANNA FREUD. New York: Schocken Books. 1964. Pp. xiii + 114. 12*s.* 6*d.*

This is a book which is a pleasure to read because of its straightforward style, its brevity and clarity, and because of its great insight into childhood problems and sympathetic understanding of them. It consists of three parts: first, a group of four lectures delivered in Vienna in 1926 on the technique of psychoanalysis of children; secondly, a paper written for the Tenth International Psycho-Analytical Congress of 1927, about the theory of the analysis of children; and thirdly, 'Indications for Child Analysis', written for the *Psychoanalytic Study of the Child* in 1945. These various lectures and papers form a useful brief exposition of the problems of child neurosis and its treatment by analysis, and it is an advantage to have them in the form of a paperback.

The author explains clearly her problem that children come to analysis with quite a different attitude from that of the adult neurotics, for whom psychoanalysis was designed. They do not accept the fact or implications of their illness, and they are not responsible for undertaking the treatment. Consequently, according to Miss Freud, their approach to the fundamental technique of free-association is different. The adult foregoes his normal control of thoughts for the purposes of analysis, under the influence of a wish to recover from a recognized neurotic disability, and a voluntary acceptance of the conditions of analysis. The child does not, and according to her this makes a fundamental difference. She rejects Melanie Klein's attempt to replace ordinary adult free-association by free play with toys by the child, and she proposes that children should be approached in a preliminary way which will secure their understanding and acceptance of the conditions of analysis. In certain cases described she shows how this was done. Her interpretations of the cases, and the mechanisms of neurosis in them, are clarity itself and leave nothing to be desired. Similarly, her theoretical explanations in the later portion of the book are a pleasure to read.

However, to the reviewer it does not seem convincing that the steps she described to bring a child analysis within the same framework as a conventional adult analysis would be either necessary or desirable in every case, and perhaps they would not always be practicable. It is also not convincing that, as she suggests, a child might bring thoughts and ideas to the analysis solely because he had encountered certain objects or situations just before he came, and not because they also had emotional meanings for him as unconscious determinants. The reviewer does feel it quite convincing that free play with toys, whether by adults or children, is revealing of unconscious forces and stress just like dreams. Everything the child brings to the session seems to the reviewer to be meaningful.

Even if every aspect of the book does not excite wholehearted agreement, it is a work which can be read easily and with profit by students and professional workers alike. One can look through it thoughtfully and find something illuminating and rewarding on nearly every page.

R. W. PICKFORD

The Development of Motives and Values in the Child. By LEONARD BERKOWITZ. New York and London: Basic Books. 1964. Pp. viii + 114. 18*s*.

The growing interest in the United States in the work of Jean Piaget has resulted in the publication of a number of studies covering a wide variety of aspects of child development. This interest is no doubt reflected in the choice of the second of the two areas selected for discussion in this volume, the moral development of the child, which opens with a presentation of Piaget's work.

He expands the section to cover the concept of conscience and resistance to temptation, and brings in the results of studies from many sources to enrich his review of the development of moral judgement and moral behaviour, thus enabling us to appraise the value of Piaget's work by seeing it in perspective.

The author has selected for treatment only two aspects of child socialization, the achievement motive and moral growth. The book fills a gap particularly perhaps in treating the first process. It is intended, like others in this series, to help the teacher (and student) of psychology, and 'to satisfy the intellectual curiosity of intelligent laymen . . . who are keeping an eye on the advance of civilization' and who will read it to understand what psychologists think and know.

It is not an easy task to write for two different types of readers, but the author, while presenting a useful overall review, has kept his style simple, and has presented facts and issues in a simple and clear fashion. He has assembled together and integrated the results of a considerable body of research in both fields, and child psychologists will find the book a most useful guide for their more detailed study. He has gathered together a good deal of material relating to the various home and cultural influences which help to shape the development of values and attitudes in the child and his references will be consulted alike by teachers and students.

Parents will find his section on changing psychological approaches to the treatment of the child of interest, but may find his review rather more difficult to assimilate, unless, like psychologists, they already have some notion of the concepts of motive, value and morality. In summary this is a useful and most readable book whose appearance has been needed for some time.

GERTRUDE KEIR

Social Organization and Behavior: a Reader in General Sociology. Edited by R. L. SIMPSON and I. H. SIMPSON. New York: John Wiley. 1964. Pp. xii + 457. 40*s*.

The student of sociology (like the practitioners of allied disciplines) can easily be confused by the variety of its subject-matter. He may also be confused by the different traditions which have contributed to the formation of the subject, from the pragmatic concerns of Anglo-American reformers to German historiography. In terms of sheer labour he is expected to have some acquaintance with the major classics, to grasp the methodological problems of the cultural sciences and the philosophical issues associated with them, to comprehend the relation between a somewhat slippery conceptual apparatus and data, and to absorb a large amount of fairly specific material which might range from colour prejudice in Liverpool to the position of eunuchs in the power structure of Byzantium. The proliferation of readers and introductions is therefore hardly surprising.

However, readers have to be based on firm criteria. This particular example is an attempt to reduce the specific material of sociology to manageable proportions, and the result is a heavy concentration on material about the contemporary United States. The section on race is exclusively concerned with America rather than (say) Liverpool or Ceylon; the section on government and politics is not much bothered by comparative material from Byzantium. In other words, comparisons across cultures and across time almost disappear from view, except in the sections on 'The Emergence of Modern Society', 'The Family' and 'Contemporary World Change'. The other concerns of sociology mentioned in the paragraph above are also excluded. There is no material from the classical sociologists and therefore no reference to large scale theoretical frameworks. There is no hint of the variety of approach found in different intellectual traditions. There is no explicit discussion of conceptual problems, although one gathers this is to be found in an accompanying text-book.

The points just mentioned are not critical but simply indicate the scope and purpose of the book. Within these limits it is useful and discriminating. The basic theme is social organization and, as the Preface states, 'values, cultural norms, socialization, interpersonal influence and reference group behaviour...are in fact bound up with the general processes of social organization'. The introductions to each section (e.g. Education, Population, Stratification, Race Bureaucracy, Social Control, etc.) are particularly useful in briefly providing a wider context and a minimum conceptual basis, together with references and footnotes indicating important contemporary or classical discussions. Moreover, there is a large number of distinguished contributors and what they provide is genuinely sociological, i.e. not information or commentaries on social statistics or social history, but controlled insights into structures and processes.

DAVID MARTIN

The Industrial Rehabilitation of Long-Stay Schizophrenic Patients. By J. K. WING, D. H. BENNETT and JOHN DENHAM. Medical Research Council Memorandum, no. 42. London: H.M.S.O. 1964. Pp. 42. 4*s*. 6*d*.

The authors report the results of an important follow-up study carried out with male chronic schizophrenic patients who were selected by an objectified standard clinical interview for industrial re-training at a Ministry of Labour Industrial Rehabilitation Unit. The investigators found that 53 % of their sample of 'moderately handicapped' patients were in employment one year after training, but almost one half of these were in sheltered workshops. Over one third were living in the open community, but less than half of these under conditions of apparent normality. When employable in sheltered workshops, the schizophrenics showed no greater handicaps than physically handicapped men with a similar training.

This study, and the conclusions drawn from it, might have gained in weight if objective multiple psychological assessments had been carried out at the same time to test for specific behavioural responsiveness to training. Other published work would have enabled the authors to predict, for example, that degree of residual intellectual capacity as well as evidence of improvement on intelligence tests following training, are relevant variables.

Response to treatment of, say, anaemia is no longer judged by external examination of the patient, and it may be felt that the undiminished flow of schizophrenic disorder deserves all the objective skills that science can offer. It is, therefore, a pity that these investigators, in an official report which is likely to be read by many people, do not visualize a clear role for psychologists in procedures which should include the repeated assessments of intellectual, general cognitive, vocational, social and emotional adaptive skills, the matching of industrial activities to available skill levels and therapeutic support while higher skill levels are being acquired. In the absence of this conception of a fully rationalized, objective and progressive procedure for reaching optimal rehabilitation levels in patients with psychological deficits, it may be felt that this report contains less a programme than the hope of a programme for this particular sociopsychological problem.

VERNON HAMILTON

A Hundred Years of Psychology. By J. C. FLUGEL. Third edition revised by D. J. WEST. London: Gerald Duckworth. 1964. Pp. viii + 394. 30*s*. London: University Paperbacks, Methuen. 1964. Pp. viii + 394. 21*s*.

This book was first published in 1933. A second edition, published in 1951, contained the first edition unaltered, with an additional section reviewing developments from 1933 to 1947. This third edition retains Flugel's original text from the first issue of 1933 but replaces the section covering 1933 to 1947 by a new part of 52 pages written by D. J. West for the period 1933 to 1963. There is an additional bibliography to go with West's new section and the chronological table has been brought up to 1962.

The main part of this book, then, has been previously reviewed in this *Journal* (1934, **24**, 450) and the anonymous reviewer found it to be an attractive and stimulating work covering the chosen areas with remarkable accuracy. However, I cannot agree with the reviewer when he suggested that the book would be useful for the general reader or the beginning student. I would regard it as a valuable second text or review for those students who have already done some

reading in the history of the subject. Flugel's writing has a very attractive wit and ease about it. His knowledge of out-of-the-way facts was astonishing at times, but the book is chiefly useful in the way he managed so neatly to integrate the separate strands of Psychology's uneven development. Read as a whole, it will be a valuable clarifier to students who are confused by the necessarily more detailed and larger alternatives available.

D. J. West, who has added the section covering 1933 to 1963, is a psychiatrist of broad interests. He has studied and written in the fields of extrasensory perception, homosexuality and criminology. The part he has had to do is a difficult one and he has chosen probably the only means of covering thirty years in just over fifty pages: '...one can do no more than list some outstanding trends and refer to a few researches chosen almost haphazard according to one's particular interets, prejudices and limitations.' It is unfortunate, though, that this final part runs as a continuous narrative. A break-down into the areas covered under separate headings would have made it easier to read.

The editing of this additional section has been very poor indeed. Some names mentioned are not in the index, and no reason is given for this. Books are mentioned by author and date but sometimes they are not in the bibliography to this new section. Hull did not formulate the Law of Effect and Tolman's classic paper is not entitled 'There is more than one King of Learning'.

BERNARD R. SINGER

Thinking: From Association to Gestalt. By JEAN MANDLER and GEORGE MANDLER. New York: John Wiley and Sons. 1964. Pp. x+300. 38*s*.

This is a collection of readings in the history of the psychology of thinking, each of which is linked by a commentary by Jean and George Mandler. Each extract is placed in its context and the development of a specific tradition in cognitive psychology is traced in the background to the selected readings. The book has two great merits. First, the commentaries although brief are scholarly and helpful in making the selections intelligible; e.g. the bits from Marbe and Ach are followed by an account of Bühler's work and Wundt's attack on it which leads neatly into the extract from Tichener. Secondly, this book gives a dozen extracts from German sources, seven of which have been translated into English for the first time by the Mandlers, and others of which are not easily accessible: Mayer, Orth, Marbe, Messer, Ach, Watt, Kulpe, G. E. Muller and Selz are all represented as well as some early Koffka. The commentaries supplying the background to these sections are especially interesting. In other respects the selection might be criticized. Just under half of the whole book is given to extracts from Aristotle, Hobbes, Locke, Hume, Hartley, James Mill and Bain. These ancients are readily available and have plenty of scholarly commentaries. It might have been more useful to have included Dilthey and Husserl; Binet and Claparède and to have found room for a representative of Behaviorism; and what of William James and Dewey? Perhaps it is time for the history of psychology to be treated more lavishly. Even if Humphrey, Bruner, Bartlett, Piaget and the cyberneticians are considered too recent, at least some of Duncker's contemporaries might have been included. In spite of these qualifications this book is a useful addition to that rare and possibly curious category of source books for the history of psychology.

ROBERT THOMSON

L'Organisation Perceptive: Son Rôle dans l'Évolution des Illusions Optico-Géométriques. By E. VURPILLOT. Paris: Librairie Philosophique J. Vrin. 1963. Pp. 187.

The visual illusions, which aroused the curiosity of Helmholtz, Hering and many others during the nineteenth century, have never ceased to interest psychologists, because these illusions present particularly striking examples of percepts which resolutely refuse to conform to the physical properties of the stimulus material. There are of course numerous situations in which perception is non-veridical. But the visual illusions are some of the most obvious and readily measurable (especially illusions involving the comparison between the magnitudes of different parts of the figures); and they are some of the least easy to explain satisfactorily. In recent years, they have been extensively investigated by Piaget and his collaborators, especially their variations with age changes in children. These observations have played an important part

in Piaget's theories as to the nature of perception, particularly as regards the function of 'perceptual activity'; observations and theory are described in Piaget's *Les Mécanismes Perceptifs*.

But certain aspects of the visual illusions and the conditions which affect their magnitude required further controlled study, and this Mme Vurpillot has performed most carefully and systematically, using the best-known illusions and measuring their magnitude in relation to variations in form, number and duration of presentations, and age of observer (five years to adults). Her hypotheses as to the basis of the illusions agree with those of Piaget, namely, that the perception of these figures evolves through three stages which can roughly be summarized as: (1) global perception without attention to detail; (2) analysis of figures and some perception of their parts; (3) reorganization of the parts into new structural wholes. These stages occur at somewhat different ages for different illusions. With many of the simpler figures (Müller-Lyer, Delboeuf), the succession of stages is paralleled by a decrease in magnitude of the illusion, which is also produced by modifications in the form of the figures and by increases in time of observation which maximize analysis and re-structuration. But with the Oppel illusion, increased analytic observation tends to result in overemphasis of the parts of the 'cross-hatched' line, and hence in an increase in the illusion. In almost all cases, magnitude decreases when the figures form part of representational pictures; which seems to suggest that greater 'reality'—or perhaps greater redundancy—here improves veridicality of perception, or else that structural details are simply overlooked.

We must congratulate Mme Vurpillot for conducting her experiments with such systematic and careful control (the more so since Piaget has sometimes been accused of failure in these respects). Her exposition is extremely clear, though occasionally perhaps a little repetitive. One does regret, however, that she was unable to relate her observations to the most recent hypothesis as to the nature of visual illusions, the three-dimensional effects postulated by Gregory and others. It would seem very difficult to explain how three-dimensional effects could account for the variations in magnitude occurring in all the conditions studied by Mme Vurpillot, at least as satisfactorily as do her hypotheses (though the Poggendorf's illusion still remains somewhat of a mystery). But one must await the attempt.

M. D. VERNON

Flavour, Taste and the Psychology of Smell. By WARREN GORMAN. Springfield, Illinois: Charles C. Thomas. 1964. Pp. viii + 106. \$5·50.

This is a recent addition to the familiar series of monographs on special topics published by Charles C. Thomas. The approach is necessarily selective, its special feature being a unique summary of psychoanalytic observations and interpretations of the sense of smell. It is impossible not to be influenced by one's own attitudes to such an approach, but having declared his own inherent scepticism the reviewer may now proceed.

An interesting historical introduction to the subject is provided, with some cross-cultural observations, and a note is included on certain market research findings about the role of visual associations in flavour perception. The brief outline of the structure and functions of the special senses concerned occupies only eleven pages and is highly selective and rather uncritical, especially in view of the recent appearance of a number of authoritative compendia which have extended and revised our basic knowledge.

One chapter has the title 'Theories of evolution and the development of biological humility'. This is primarily a survey of various facets of Darwinian thought. The final chapter summarizes 'an evolutionary and psychological theory of smell'. Whereas this includes a brief account of physico-chemical theories of the characteristics of the peripheral stages of odour perception, the main emphasis is given to (Freudian) psychoanalytical observations and interpretations. It is suggested, for example, that repression has taken place at both a psychological and an organic level. The psychological processes are described as 'similar to physiologic adaptation or inhibition', whereas 'organic repression' refers to evolutionary changes which have led to changes in man's olfactory structures associated with a decrease in their effectiveness.

If the reader is seeking this particular sort of information or a stimulating treatment of the whole subject then Warren Gorman's book is a useful choice. For an outline of the more conventional, scientific background, however, much better sources are available.

ROLAND HARPER

Dream Interpretation: A New Approach. By T. M. French and Erika Fromm. London and New York: Basic Books. 1964. Pp. 224. 30*s*.

It is difficult to know whom this book will serve: in many ways it is a good one, but Dream Interpretation is a subject with a limited audience in this country. In much psychoanalytic practice dream interpretation has been subordinated and sometimes even displaced by transference interpretation and some so-called analytic forms of treatment are carried on largely at a dream interpretation level, using symbolic content interpretations as opposed to process ones.

The strength of this book lies in its description of a careful checking and re-checking of the interpretive processes. All forms of content interpretation could lead to wild speculation and the authors are at pains to prevent this. One of their principles is that they regard interpretations as hypotheses to be tested, and it is one which cannot be supported enough. Nevertheless, their method introduces a weakness. The dream material, cogently used here, was that of a patient of another analyst. Although it is obvious from his notes that he is aware of the personal relationship going on between his patient and him, this method does mean that the authors give less weight than they probably otherwise would have done to the dream thought process as a facet of the treatment and its conflicts. The inter-personal adaptations worked out in the transference relationship with the analyst need the same careful checking but are perhaps less obvious to outside observers who could be aware of the verbal content chiefly. Interpretations of this order, and therapeutically important, occur to one while reading this account but these in turn need the same sort of validation.

R. E. D. MARKILLIE

Clinical Psychology. By J. B. Rotter. Englewood Cliffs, N.J.: Prentice-Hall. 1964. Pp. xv + 112.

Introduction to Psychology for Medical Students. By R. R. Hetherington, D. H. Miller and J. G. Neville. London: Heinemann. 1964. Pp. ix + 212. 25*s*.

Considering these two books together makes one appreciate the problems of publishers. On the one hand they have to find readers to buy a good book, and on the other they have to find a good book for a ravening hoard of text-hungry medical students.

Rotter's book is admirable. It gives a clear account of truly current views on the assessment of intelligence and personality and it includes a useful abstract of contrasting psychotherapeutic systems. The whole is neatly set in historical perspective. However, the brevity of the treatment makes it only suitable for drawing the attention of students to the existence of the field. Unfortunately it cannot be recommended for this purpose in this country because it presents an American rather than a British stereotype of clinical psychology. The clinical psychologist it portrays seems to be a psychiatrist who gives tests instead of drugs. The British roles are more distinct, and our attitude to training and qualifications is different. Two omissions are worth noting. The author appears ignorant of the chromosomal abnormalities found in mongolism and there is no mention of behaviour therapy. The book is an attractive paperback but several misprints have crept in (the sole mention of language disorders comes out as the 'asphasias').

There are few texts suitable for the increasing numbers of medical students taking psychology in their pre-clinical years, but the one produced by Hetherington, Miller and Neville is hardly a breakthrough. It is certainly aimed unerringly at the medical student and he will probably find it as pleasant a nostrum as Knight and Knight. If he takes it to heart, it will probably make him a better physician, but it is less certain that it will teach him much psychology. All introductory texts need to simplify: this can be done by integrating and articulating the subject to give the student a framework which will help him develop his understanding. This text simply abbreviates. It is poorly organized: topics are dealt with patchily throughout the book with many cross-references which lead one to hope that the next chapter will really tackle the point. Some sections may go over the heads of the students, not because of the complexity of the psychology, but because they will not yet know enough clinical medicine. The style does not clearly distinguish reports of research from pedagogical paradigms and the book would be improved by more detailed references to studies quoted in the text. The basic psychology and the metaphysics are out of fashion: 'A timid person walking down a dark lane at night will tend to

see every dark shape as a robber lurking to attack. This is a muddle between the fearful image and the innocuous percept' (p. 19). Most of the section on perception is taken up with illusions. We are told that the terms cognition, affection and conation 'will be used throughout the book' (p. 3): fortunately this does not happen. Pleasure seems to be given explanatory force: 'When he is hungry again he tends to repeat his lever pressing because it gave him pleasure before' (p. 68). If we make the legitimate translation into 'because he liked doing it', the explanation loses its gloss. This approach contrasts oddly with references to recent work like that on cannibalistic planarians. Some sections show poor editing: e.g. p. 31, 'He cannot walk before he is mature enough to do so, neither will he walk unless he is taught. This process is known as *maturation*...'. Discussing Pavlovian conditioning on page 32 the sight of food is included in the unconditioned stimulus, but on page 33 salivation to the sight of food is given as an example of a conditioned response.

To balance these faults the book has brief chapters on interpersonal relationships in hospital, pain, sleep, disorientation, development and old age, and on reactions to physical illness and to drugs. This is as worthwhile as it is unusual, but medical students need some solid basic psychology under their belts as well.

P. R. F. CLARKE

Statistics for Psychologists. By WILLIAM L. HAYS. New York: Holt, Rinehart and Winston. 1963. Pp. xvi + 719. 84*s*.

There are now available a large number of statistical texts written specifically for psychologists and social scientists. Prof. Hays's volume is something more than a mere addition to the list. It is addressed primarily to the advanced undergraduate or beginning graduate student who wishes to 'go beyond the routine methods and gain more understanding of their underlying rationale'.

After a somewhat unconventional beginning with an elementary introduction to set theory leading on to probability and the binomial theorem, the book covers descriptive statistics and all the inferential techniques commonly encountered in psychology. In addition, there are sections on Bayes's theorem, Tchebycheff's inequality, central limit theorem, and the relationships between F, t, and χ^2. All this is accomplished in considerable depth using little more than elementary algebra. The reader who is discouraged by mathematical symbols need only turn to the two appendices to find clear expositions of the use of summation signs and subscripts, and the concept of mathematical expectation.

A major theme running throughout the text is what the author sees as 'the experimenter's tendency to confuse statistical significance with the importance of results for prediction'. Thus much emphasis is placed on confidence limits and measures of 'variance accounted for'. In an unusual section entitled 'Can a sample be too large?' we are warned that small and possibly unimportant differences may show up as highly significant where N is large. The point, although an obvious one, is often overlooked in practice. A useful discussion is provided of the inferences which may or may not be drawn from fixed effects and random effects models in analysis of variance. Simple experimental designs are favoured and the difficulties of interpreting a number of F ratios obtained from a complex multivariate design are emphasized.

All experimental problems and data used for purposes of illustration appear to have been artificially constructed. Sometimes this is made clear, but on several occasions words such as 'A study was made of...' are used with no indication whether or not the data given are real. At one point we are asked to entertain the hypothesis that the performance of bright children on a task may improve under the stress of competition, while that of dull children may deteriorate. A consequence of the hypothesis is that two groups of children randomly selected and subjected to differing amounts of competitive stress should differ in the variance of their scores. This example is given to illustrate a possible use of the F ratio, with no indication that the experimental design is, at best, unrealistic. There is a confusion on page 609 where a 3×4 contingency table has every expected frequency less than 4 and an incorrect χ^2 value has been calculated.

The general orientation of the book is theoretical rather than applied. It should be of great value to the kind of reader the author has in mind, and also to the research worker whose statistical knowledge needs refreshing. It can be relied upon to provide satisfactory answers to

most of the questions encountered from time to time by the user of statistics in psychology and glossed over or omitted in other works. The style is eminently readable and the text contains much incidental statistical wisdom.

J. K. CLARKSON

Artistic Self-Expression in Mental Disease. By J. H. PLOKKER. London: Charles Skilton Ltd.; The Hague, Mouton and Co. 1964. Pp. 224. 84*s*.

Artistic Self-Expression in Mental Disease is a beautifully produced book, well-printed and with ninety-five illustrations, thirty-nine of them in colour. It is a pleasure to possess this book. Prof. Plokker's text is fascinating and thought-provoking, but in the final analysis he fails to lighten the mystery of either schizophrenia or of art. He himself is, of course, fully aware that such a task is practically impossible.

For the Anglo-Saxon reader there is the added difficulty, to which Prof. Plokker draws attention when he points out that the European's approach to schizophrenia, with which he himself identifies, differs markedly from the approach to this disease in Britain and America.

Prof. Plokker has pointed out that the schizophrenic symptoms—and therewith also their pictorial productions—have recently altered, partly as the result of chemotherapy, but partly because of the change of attitude to him on the part of the hospital staff, as well as of friends and sometimes family. It is indeed remarkable now 'un-mad' many patients can become if one addresses onself to the healthy and responsible part inside them. Furthermore as the so-called normal person is becoming progressively more aware of and related to his own deep and unconscious impulses and phantasies, the gulf between the normal and the mentally sick seems to narrow.

The identification of schizophrenic art would present no difficulty at all if art movements had remained classical, or at least had never moved beyond the realist or the impressionist schools. But in a world in which gallery and museum walls display pictures in the surrealist mode, the expressionist mode, abstracts, action painting, cubism, *objet-trouvés*, collages, pop art, primitives, etc., only a very deep and sophisticated study can help us learn to differentiate—if it is possible—the products of the 'sane' from the products of the 'insane' artist.

Plokker believes that such a study demands the analysis of form rather than the analysis of content; this point seems most vital and important; he suggests, very rightly that men do not differ much one from the other on the basis of their deeply unconscious drives and phantasies, but that the difference depends on the individual consciousnesses that they develop. He is justified in his conclusion that those psychoanalysts who have concerned themselves with the problem of art but who have confined themselves only to a content analysis, have not in fact advanced our understanding of aesthetics and have consequently failed to help us distinguish works of art from the products of what has remained art therapy.

In this book Plokker attempts to fill the gap; he tries to provide some of the formal 'signs' which he considers characteristic of schizophrenic paintings, such as: (*a*) the need to fill the whole sheet; this is an expression of what he also calls 'exhuberant opulence'; (*b*) the passion to decorate, often with minute motifs, though there may be a sudden introduction of realistic motifs; (*c*) the distortion of space, by altering, for instance, the relative proportions of foreground and background; (*d*) perseverations and stereotypes; (*e*) X-ray drawings; (*f*) scenes that seem to be composed of empty, immeasurable desert-like space; (*g*) movements that give the appearance of being stiff, taut and inflexible.

Any study of art must inevitably concern itself with the problem of symbols; Plokker describes the symbol most effectively as 'a tentative description of something that is not completely clear' (p. 93). He seems to consider that Jung's approach to symbols is very much more useful to an understanding of art than is Freud's. For Freud symbols are merely defensive ploys, disguises of some few basic objects of instinctive desire; Jung ascribes to the symbol a much more comprehensive function and concerns himself also with the fact that it appeals to our whole being. I wonder if the study and comparison of normal and psychotic art might be furthered if use were made of the distinction, developed by Kleinian psychoanalysts such as Hannah Segal, between symbolic equivalence and true symbolization, the latter involving the 'symbolic attitude'.

ROSEMARY GORDON

Motivation: Theory and Research. By C. N. Cofer and M. H. Appley. New York: Wiley. 1964. Pp. 958. 94*s*.

As a curtain-raiser Cofer and Appley outline the differences which exist in the various approaches to the subject-matter of motivation and their dependence upon tradition, allegiance and interest. This outline is continued in a following exposition of historical perspectives, e.g. the emergence in the nineteenth and twentieth centuries of the distinction between 'conceptions of motivation' and 'motivational constructs', from the influence of Darwinian evolution and adaptation theories including the survival value of 'adaptation' to the postulation of characteristic action patterns or instincts as their mechanisms. In two excellent chapters on the bodily concomitants of motives, the authors reverse the traditional motivational priorities, making hunger, thirst and pain avoidance follow hoarding, sucking, migration, temperature regulation, sleep, inhibition (fatigue), as well as sexual and maternal behaviours, on the ground that these latter forms of conduct are not only adaptive but cyclic and largely independent of experience, whereas the former are somewhat artificial and dependent upon deprivation and experimental manipulation.

Theories of motivation, for many centuries, have rested on the notion of a fundamental natural balancing tendency which has been known since the time of Cannon as 'homeostasis'. In an outstandingly clear exposition and discussion the authors emphasize the view that dynamic equilibrium is not only an organismic regulating mechanism which compensates in the face of disturbance, but also, at the behavioural level, the essential mechanism of evolutionary development, in spite of the charge of non-applicability to some of the more complex and altruistic of human motives, and in spite of the clearly mechanistic and, not so clearly, vitalistic forms it takes.

The middle chapters of the book are concerned with a series of topics both too detailed and too heavily documented to be dealt with in a review. Notable among them are the treatments of hedonistic and activation theories of emotion as motive, of conflict, stress and frustration as 'drives' and of the place of motivation in Learning Theory.

The later parts of this work, although easier to read, are sometimes harder to understand. In them the authors present chapters on psychoanalytic motivation theory, self actualization and social motivation. The first and third are differentiated largely, but not exclusively, by the emphasis placed by their exponents on the internal and external factors initiating conduct. The self actualization concept stresses the nature of 'human nature' and is thus a metaphysical but possibly useful device for explaining if not describing behaviour.

After a belated 'declaration of intent' (a sympathetic exposition of existing theories and their shortcomings) Cofer and Appley present their own 'Unified theory of motivation'. They postulate two mechanisms based upon the universality of an equilibrium model and the anticipatory aspects of consumatory behaviour. The first they call an 'anticipation-invigoration mechanism' (A.I.M.) and the second a 'sensitization-invigoration mechanism' (S.I.M.). Both lead to augmentation of vigour and both are based upon deprivation. However, whereas the former stresses the anticipations developed by deprivations associated with stimuli under conditions of arousal, the latter emphasizes an unlearned deprivation-arousal-sensitization process. Conjointly these mechanisms are advanced by the authors to offer a more useful concept than 'drive', which they believe to be 'without utility'. In brief they would supplant 'drive' by A.I.M. and S.I.M. both of which are mechanisms of invigoration which offer more fruitful reconceptualizations of motivation. They are more fruitful because these concepts will compel researchers to enquire into different criteria within a two dimensional framework rather than adding supplementary concepts to a single undifferentiated index. We have in fact, a two factor theory of motivation. The reader must make his own choice as to whether A.I.M. and S.I.M. are more than semantically differentiated from learned and unlearned drives.

H. C. HOLLAND

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology

The British Journal of Psychology (Monograph Supplements)

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology (about 512 crown quarto pages per volume), *The British Journal of Medical Psychology* (about 350 crown quarto pages per volume), and the *British Journal of Social and Clinical Psychology* (about 320 crown quarto pages per volume) appear in quarterly parts, the subscription price of each, payable in advance, being £4. 10*s.* per volume post-free ($16.50 in the U.S.A.). Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, N.W. 1. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 288 pages. The subscription rate, payable in advance, is £3. 10*s.* per volume post-free ($13 in the U.S.A.). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, London, E.C. 4; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number. The subscription price in the U.S.A. is $4.20. Subscriptions should be sent to the Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36 Essex Street, Strand, London, W.C. 2.

A series of *Monograph Supplements* is issued in connexion with *The British Journal of Psychology* but is not included in the subscription for that journal. Orders may be sent to any bookseller, or the Cambridge University Press, 200 Euston Road, London, N.W. 1, from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor A. Summerfield, Department of Psychology, Birkbeck College, Malet Street, London, W.C. 1;

for *The British Journal of Medical Psychology* should be sent to

Dr T. Freeman, Lansdowne Clinic, 4 Royal Crescent, Kelvingrove, Glasgow, C. 3;

for *The British Journal of Social and Clinical Psychology* should be sent to

Mr Michael Argyle (Social Psychology Editor), Institute of Experimental Psychology, 1 South Parks Road, Oxford, *or* to

Dr R. R. Hetherington (Clinical Psychology Editor), 6 Abercromby Square, Liverpool 7;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Dr R. J. Audley, Department of Psychology, University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Mr L. B. Birch, The Institute of Education, The University, Sheffield 10.

for *The Monograph Supplements* should be sent to

Professor I. M. L. Hunter, Department of Psychology, The University, Keele, Staffs., whose decision regarding the publication of all Monograph Supplements will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1 September of the preceding year.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £6. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1 annually), *The British Journal of Medical Psychology* (£1 annually), *The British Journal of Social and Clinical Psychology* (£1 annually), *The British Journal of Mathematical and Statistical Psychology* (£1. 10*s.* annually), and *The British Journal of Educational Psychology* (£1 annually).

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, Tavistock House South, Tavistock Square, London, W.C. 1.

Preparation of Papers for the British Journal of Psychology

In submitting papeis for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.
It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references, and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. 56 (1965) onwards.
A brief summary, suitable for use as an abstract, should precede the article.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.
Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.
The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.
Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–48.

GALTON, F. (1883). *Inquries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *L'Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.
The Oxford Dictionary is taken as the authority on spelling.
For other general guidance, authors are referred to *The Journal of Physiology: Suggestions to Authors*. Cambridge University Press. Price 2*s*.

THE BRITISH JOURNAL OF PSYCHIATRY

Clinical:

Pathologic Jealousy and Psychochemotherapy; by HORACE B. MOONEY.

Neurotics Who Fail to Take Their Drugs; by RONALD S. LIPMAN, KARL RICKELS, EBERHARD H. UHLENHUTH, LEE C. PARK and SEYMOUR FISHER.

The Outcome of Obsessional Disorder. A Follow-up Study of 100 Cases; by LINTON GRIMSHAW.

Social Psychiatry:

Factors Underlying the Retention in the Community of Chronic Schizophrenics; by L. S. GILLIS and M. KEET.

Ethnopsychiatry in Central Australia. I. 'Traditional' Illnesses in the Eastern Aranda People; by J. E. CAWTE.

Ethnopsychiatry in Central Australia. II. The Evolution of Illness in a Walbiri Lineage; by J. E. CAWTE and M. A. KIDSON.

Forensic Psychiatry:

Murder in the West of Scotland; by HUNTER GILLIES.

Drug Treatment:

A Controlled Trial of Mebanazine ('Actomol') in Depression; by J. C. BARKER, I. A. JAN and M. DAVID ENOCH.

A Comparison of Nortriptyline and Amitriptyline in Depression; by JOHN T. ROSE, M. R. LEAHY, IAN C. A. MARTIN and T. T. WESTHEAD.

A Controlled Trial with Trimipramine, a New Anti-Depressant Drug; by M. M. SALZMANN.

Acute Hepatic Necrosis Following the Treatment with Amitriptyline and Diazepam; by M. L. CUNNINGHAM.

A Point of View:

Cryptomnesia and Plagiarism; by F. KRAUPL TAYLOR.

Correspondence.

Reviews.

The Annual Subscription for Volume III (1965) is £6 (U.S. $16.80) post free.
Orders and remittances should be sent direct to

H. K. LEWIS & CO., LTD.

136 GOWER STREET, P.O. BOX 66, LONDON W.C. 1.

or through any Bookseller.

Brit. J. Psychol. (1965), **56**, 4, *pp.* 349–358
Printed in Great Britain

SEARCHING THROUGH WORD LISTS*

BY ULRIC NEISSER AND HENRY K. BELLER
Brandeis University, Waltham, Mass.

Two experiments are reported in which subjects searched through lists of words, looking for targets defined only in terms of their meaning (e.g. 'any animal'). Under these conditions, scanning is much slower than when the target is a known word, or a member of a small fixed set. It is argued that printed words are processed in two stages, termed 'stimulus examination' and 'memory examination' respectively. Preliminary data relevant to other aspects of the scanning method are also presented: error rates, the importance of motivation, the extent to which words scanned over can be recognized later, and the relation between scanning and reading aloud.

What are the cognitive processes that begin when a printed word falls on the eye? Though they are complex, at least the early stages of processing may be open to experimental analysis.

Let us assume that at some particular point a cognitive subsystem is aroused which corresponds to the linguistic unit—the word—that is seen. Whether this subsystem is regarded as a cell-assembly, a group of traces, a subroutine, or in some other way is irrelevant for the present argument. Its activity amounts to a tentative decision by the perceiver about what the word is. To be sure, this decision is neither the beginning nor the end of the processes initiated by the stimulus. The events that precede it are, however, of a different character from those which follow. This critical point may be called the 'identification' of the word. Identification is itself not a process, but the termination of one kind of activity and the beginning of another.

The processes which precede identification may be called 'stimulus examination'. Perhaps it would be appropriate to call them 'perceptual' as opposed to 'intellectual', but we prefer to avoid these terms. Stimulus examination ensures the correct identification of words over a wide range of type-styles and handwritings; in spite of omissions and distortions of letters; in a variety of different contexts; and independently of retinal position or size. To accomplish this, the activity in the receptors is analysed for the possible presence of a variety of 'features' which are more or less invariant over such transformations, and information about these features is somehow combined to arrive at the identification of the word. The features need not be individual letters; in some cases words may be identified on the basis of characteristics such as length and shape.

Our previous studies of visual search have dealt chiefly with this process of stimulus examination. The targets were not words but single characters, so the critical features must have been the figural properties of individual printed letters. These studies suggest that stimulus examination can be terminated before the point of identification is reached. In looking through a list for the letter 'K', subjects do not 'see' the irrelevant letters through which they search. In other words, these letters, though

* Much of this work was performed while the senior author was at Lincoln Laboratory of the Massachusetts Institute of Technology, which operated with support from the U.S. Air Force. In addition, the research was partially supported by the National Science Foundation under grant no. G-21654 to Brandeis University.

examined, are not identified. Identification does not occur until the successful termination of the search. Moreover, the scanning rate is independent of the number of targets, or potential identification points (Neisser, Novick & Lazar, 1963), though it may vary with the nature of the irrelevant letters (Neisser, 1963, expt. III). Thus there seem to be two phases of stimulus examination: an initial process of 'feature-formation' is succeeded by one of 'letter-formation', which ends when the letter is identified. The results of the earlier studies are, then, (*a*) that inspection may be terminated after feature-formation and before identification, and (*b*) that some of the processes in feature-formation are better thought of as parallel than as sequential.

The processes which follow identification of a word operate upon very different information from those which precede it. The physical dimensions of the stimulus become irrelevant; indeed, the subsequent processes can go on in the absence of any stimulus. They consist of forming and using relations between the identified word and other information previously stored. Other word-systems may be aroused by what is often called 'association'; the meanings of the identified word may be determined; visual images may form; properties of the object which the word denotes may become available, and so on. Because these activities all involve stored information, they may be called 'memory examination'. The search experiments to be reported here differ from our earlier work in that memory examination, as well as stimulus examination, was necessary to discriminate between the target and the irrelevant items. That is, some of the targets were defined by meaning rather than by their figural properties.

The subjects of the present experiments scanned through word lists in search of certain target words. The use of words as targets is not enough in itself to ensure that memory examination will be involved. In looking for a given target word, say 'MONDAY', a person can act much as if he were searching for a single letter. He can recognize the target by its visual configuration alone, forgetting about what it means. Such a search would require only stimulus examination, and should proceed very rapidly. Even an ensemble of possible targets may require only stimulus examination, if all the words in the ensemble are well established before the search begins. In one condition, the target of the search was defined as any of the fifty American states, but only the eleven with six or fewer letters actually appeared. We expected that these eleven words would soon become familiar enough to be treated simply as stimulus configurations, independently of their meaning. This expectation was not entirely confirmed, as will appear below.

To make sure that the processes of memory examination would be used, some targets were defined in terms of meaning alone. In one condition the subject had to find 'an animal'; in another, 'a proper first name'; in a third, either one of these. On encountering the word 'HOUSE', say, in such a search, one must consult stored information to determine that it does *not* refer to an animal. Perhaps it is not actually necessary to decide what 'HOUSE' does mean, but one cannot rely on stimulus examination alone. There are far too many animal words to establish their configural properties in advance. In other words, a meaning-defined target forces the subject to 'read' each word in far greater depth than does a single-word target.

The first hypothesis to be tested in the present experiment was, therefore, (1) that scanning will be faster when the target can be distinguished by stimulus examination

alone (single words, small sets of words, letters) than when its definition requires memory examination (target classes defined by their meaning). In addition, the experiment was designed to test two other, more speculative hypotheses: (2) that all searches for targets defined by stimulus examination alone will be equally fast; (3) all searches based on meaningfully defined targets will be equally fast.

These two latter hypotheses were formulated because our previous studies of letter-searches had shown that the number of targets does not affect the speed of searching. Thus, at least some processing operations can be carried out in parallel at the level of stimulus examination. If all operations were parallel and of equal complexity, then all searches at a given level would be equally fast.

Experiment I, involving six scanning conditions, was carried out to test all three hypotheses. It will appear below that hypothesis (1) was confirmed, hypothesis (2) disconfirmed, and hypothesis (3) left ambiguous. Experiment II was carried out as a further test of hypothesis (3).

Experiment I

Method

Stimulus materials. The experiment required three pools of English words, one of which contained names, another animals, and the third general words (i.e. non-names and non-animals). These were constructed with the aid of the Lorge Magazine Count (Thorndike & Lorge, 1944). Construction began with a computer tape on which 77,029 different words and their frequencies of occurrence were represented. The tape was edited by computer to eliminate words containing incorrect symbols or other detectable faults, as well as words that were too long (over eight letters) or too rare (less than three occurrences in the five million words of the magazine count). The resulting list of 24,869 words was twice edited by hand to eliminate words which seemed to be improper or dubious English, and to separate names and animal words into special pools. Further editing by computer then eliminated words which were shorter than three letters or longer than six. Words containing the letter 'K' were also set aside into a special pool.

After these editing procedures, the general pool contained 7901 words; the K-pool had 739 words; the name pool had 514 men's and women's first names; and the animal pool contained 115 animal words, broadly defined to include birds, fish, insects, etc. Another target pool contained the eleven names of American states which fit the length criteria: Alaska, Hawaii, Idaho, Iowa, Kansas, Maine, Nevada, Ohio, Oregon, Texas, Utah. Since the word 'Monday' was the target in one experimental condition, all days of the week and months were removed from the general pool, as were a few other words which had special significance in another experiment (Neisser & Stoper, 1965). All the pools were mutually exclusive, and words which might reasonably be in both the name and the general pools (e.g. 'pat') were eliminated altogether. During the course of the experiments, a few other names, animals, or ambiguous words came to light that had been overlooked during the editing process. These were eliminated from the general pool as soon as possible after they were discovered.

The lists to be scanned were constructed from these pools by a computer, and printed on individual pages. Each list was a column of fifty words, of which forty-nine were drawn at random from the general pool. The remaining item, whose position varied unpredictably from one list to the next, was the target, drawn from one of the other pools according to the experimental conditions described below. Fifteen lists pertaining to a single day's run in a single condition were attached together to form a 'deck'; this facilitated the use of the same order of presentation to all subjects. For the first 10 days, the lists were printed on partially lined paper; thereafter blank white paper was used. No discontinuity in scanning speed resulted from this change.

Despite the comparatively large size of the name and animal pools, it was inevitable that some individual words would occur repeatedly in the course of the six-week experiment, and thus subjects might improve simply by becoming familiar with specific target items. To check on

this possibility, one-eighth of the words in each pool, selected at random, were set aside. They did not appear as targets during the first 25 days, but were added to the target pools for the first time on day 26.

Experimental design. There were six experimental conditions. In condition *K*, the target was any word containing the letter K. In condition *Monday*, the target was always the fixed word 'MONDAY'. In condition *Name*, the target was drawn at random from the name pool described above. No list of all the targets was ever presented to the subject, who was simply instructed to look for men's and women's first names. In condition *Animal* the target was similarly drawn from the animal pool, with the subject instructed that 'animals' included birds, fish, insects, and the like. These four conditions were presented on every experimental day. The procedure took nearly an hour at first, but the time decreased as the subjects became more proficient. On day 11 we felt able to add a fifth condition, *States*, to the subjects' duties. The instructions were to look for any American state. No explicit list of states was presented. Search rates in these five conditions seemed to have levelled off by day 17, so a sixth condition, *Name or Animal* (*N or A*) was introduced. In *N or A* the target could be either a name or an animal so that the subject had to remain alert for both possibilities. In fact, half the targets (at random) were drawn from each pool. The order of the six conditions was varied from day to day.

The experiment continued for twenty-eight daily sessions (on a 5-day week). Thus, the subjects did not have equal practice in all the conditions. We were not concerned about this disparity, since our primary interest was in the asymptotic scanning rates—the points at which practice seemed to produce no further effect. Moreover, we assumed that positive transfer between the various conditions would partly compensate for the unequal amounts of practice.

At the end of the experiment, certain additional procedures were undertaken to find out whether the irrelevant (non-target) items were noticed and remembered. In a brief inquiry on day 28, the subject was simply asked whether he had noticed the word 'FRIDAY' which had appeared on a K-list; none had. On day 29, only conditions *Name*, *K*, and *States* were given. Immediately after the ninth, seventh, and ninth list in these conditions respectively, a recognition test of memory was administered. Each subject was read sixteen words, including eight from the list he had just scanned. He was asked to identify those which had appeared on that list.

Apparatus and procedure. The lists appeared behind a glass window in a specially constructed box. When the subject (run individually) was ready to begin a search, he activated a hand-held rotary switch which turned off a fixation light, turned on lamps illuminating the list, and also started an electric timer. He immediately began to search down through the list for the type of target appropriate for the given experimental condition. When he found it, he turned the switch again (stopping the timer) and said the target word aloud; then a third turn of the switch caused the list to disappear. After the experimenter had recorded target position and search time, he pulled the next list into position behind the window and indicated that a new search could begin. The subject was instructed to scan steadily downward, avoiding regressions. The linearity and regularity of the search-time functions indicates that these instructions were generally followed.

Twelve lists were presented daily in each condition. The first two were considered practice and were not used in the computation of scanning rate. Errors frequently occurred, either because the subject overlooked the target and reached the bottom of the list without finding it (omissions) or because he responded to a non-target word (commissions). Search times on error-trials were not recorded, but extra trials were given subsequently to make up for them so long as the total number of lists did not exceed fifteen.

Subjects. Two male and four female students at the Harvard Summer School served as subjects. They were paid by the hour. All of them knew that the experiment was an attempt to determine the maximum possible scanning rates for these materials. In all conditions, they were repeatedly encouraged to go faster.

Results

Method of analysis. The ten daily search times for each subject in a given condition were analysed to determine time per word (the inverse of scanning rate). The method of analysis has been described in previous papers (cf. Neisser, 1963). The search times are plotted as function of the position of the target in the list, a least-squares line

is fitted to them, and the slope of this line is taken to represent the time needed to examine each successive item.

Mean scanning speeds. In Fig. 1, time per word for individual days and conditions has been averaged over subjects. It is clear that *Monday* was the fastest condition from the beginning, and consistently required about 0·07 sec/word after day 18. *States* was slower than *Monday*, even after extensive practice, levelling off at about 0·11 sec. The difference was maintained consistently by all six subjects through day 26, and by four of them through day 28.

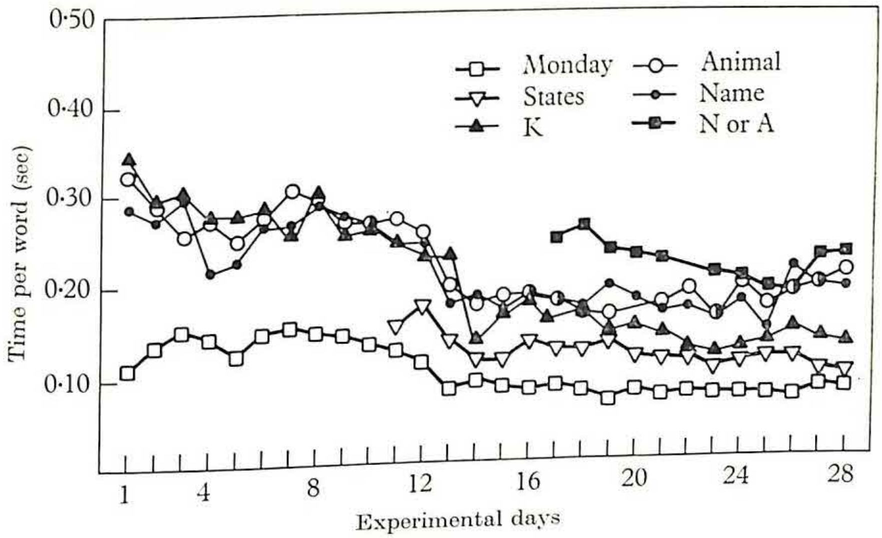


Fig. 1. Time per word as a function of day at practice, Expt. I. Each point is the mean for six subjects.

It is apparent from the figure that *K* was initially very slow, but eventually reached a speed only slightly slower than *States*, and faster than any of the meaningfully defined conditions. However, there were wide individual differences in the scanning rate for *K*, so that this outcome cannot be regarded as conclusively established.

All of the meaningfully defined conditions (*Name*, *Animal*, *N or A*) were scanned more slowly, on the average, than any of the three conditions, already mentioned, which were definable by visual inspection alone. The differences between the meaningful conditions on the one hand and *Monday*, or *States*, on the other appeared consistently for all six subjects in the later stages of the experiment, and so may be regarded as significant at the $(\frac{1}{2})^6$, or 0·02 level. Again, the wide variability in condition *K* prohibits any firm conclusions.

Fig. 1 shows that *Name* and *Animals* were equally fast, about 0·17 sec/item over days 15–25. *N or A*, in which the target might belong to either class, was scanned somewhat more slowly than the other two by most subjects. Individual data appear in Table 1.

Errors. An error of *omission* is a failure to find the target; an error of *commission* is a response to a word other than the target. The latter occur extremely rarely in *Monday* (3), *States* (7) and *K* (4). In the cases of *Monday* and *States*, nine of the errors

of commission were made to the words with great physical similarity to a target: 'MODIFY' ('MONDAY'), 'ORIGIN' (OREGON'), etc. One subject responded to 'MASS' —an interesting error to be discussed below. Errors of commission are also rare in the three meaningfully defined conditions; most of those which do occur are traceable either to a word that ought to have been edited out of the list, or to understandable misapplication of a criterion; e.g. 'CLAUSE' interpreted as a name.

Table 1. *Time per word (sec) by subject and condition, Expt. I, averaged over days 17–28*

Condition	Subject						
	SS	LL	KP	SW	GA	RD	Mean
(1) 'Monday'	0·05	0·05	0·09	0·08	0·08	0·09	0·07
(2) 'States'	0·07	0·09	0·11	0·13	0·11	0·15	0·11
(3) 'K'	0·07	0·13	0·21	0·13	0·13	0·14	0·14
Mean 1–3	0·06	0·09	0·14	0·11	0·11	0·12	0·11
(4) Name	0·13	0·14	0·15	0·19	0·19	0·24	0·17
(5) Animal	0·16	0·16	0·15	0·18	0·20	0·21	0·18
(6) N or A	0·15	0·18	0·20	0·21	0·25	0·29	0·21
Mean 4–6	0·15	0·16	0·17	0·19	0·21	0·24	0·18
Grand mean	0·11	0·13	0·15	0·15	0·16	0·19	0·15

Table 2. *Percentage errors of omission, by subject and condition, Expt. I*

Condition	Subject						
	SS	LL	KP	SW	GA	RD	Mean
(1) 'Monday'	16	18	12	4	4	5	10
(2) 'States'	19	12	9	6	5	7	10
(3) 'K'	8	20	10	8	2	6	9
Mean 1–3	14	17	10	6	4	6	10
(4) Name	22	18	28	13	15	31	21
(5) Animal	26	27	29	17	12	30	24
(6) N or A	42	35	28	16	17	38	30
Mean 4–6	30	26	28	15	16	33	25
Grand mean	22	22	19	11	9	20	17

Errors of omission, on the other hand, occur frequently. Table 2 shows how the rates at which these errors were committed varied for individual subjects in different conditions. Errors are much more common in the meaningfully defined conditions than in the others, reaching a remarkable 30 % in *N or A*. In general, the conditions which produce the slower scanning rates also give rise to the most errors. However, Tables 1 and 2 indicate that this relation does not hold among individuals: while the slowest subject makes many errors, the fastest makes even more.

As previously noted, a number of words were added to the name and animal pools on day 26, and some of these words did appear as targets in the last three days of the experiment. The error rates on these new targets were in no way different from the rates for the remaining targets (many of which had appeared on earlier days). Combining subjects and conditions for all three of these days, there were 51 instances of such new targets, eliciting 14 errors (0·27); there were 399 instances of other targets, many of which had occurred before, eliciting a total of 110 errors (0·28).

Recognition tests No subject performed better than chance on day 29 when, after several lists, he was asked to say which of sixteen words read to him was from the list he had just finished scanning. The usual report was that few, if any, of the words appeared familiar.

The results of Expt. I substantiated hypothesis (1): the search tasks which involved memory examination were performed more slowly than those requiring only stimulus examination. Hypothesis (2) was disconfirmed, in that both *States* and *K* were slower than *Monday*. An interpretation of this finding appears in the discussion section below. The data seemed ambiguous, however, with respect to hypothesis (3): while *N or A* remained slightly slower than *Name* or *Animal* throughout, it seemed to approach them as the experiment ended. Given additional practice, subjects might be able to scan with equal speed in all three of the meaningfully defined conditions. Expt. II was carried out to explore this possibility.

Experiment II

Method

Stimulus materials. Since the purpose of Expt. II was to replicate conditions *Name, Animal,* and *N or A* of Expt. I, targets were drawn from the Name and Animal pools of that experiment. For days 1–16, and also on day 20 (the last day of the experiment), the irrelevant words on the list were drawn from the same pool of general words used before, in which the minimum language-frequency was 3 per 5 million. To explore the effect of word-familiarity on scanning rate, the irrelevant items were restricted to more common words (frequencies of at least 50 per 5 million) on days 17–19.

Apparatus and procedure. The procedures of Expt. I were followed, and the same apparatus was used.

Subjects. Three male and three female Brandeis undergraduates served as paid subjects. The special encouragement given them to induce rapid scanning is described below.

Experimental design. To keep the first experimental session reasonably short, only conditions *Name* and *Animal* were presented; thereafter all three conditions were presented daily.

Systematic efforts to maximize performance were made in this experiment. On days 1–9 these efforts were restricted to praising the subjects' work after every 12–15 lists (one condition). After the ninth daily session, the experimenter took 15 min to explain that it might not be necessary to identify every word absolutely; that it would help to keep one's eyes moving rapidly down the list; and that perhaps skill in rapid visual scanning might transfer to everyday reading tasks. During days 10–14 he offered almost continuous verbal encouragement, and evaluated the subjects' performance after each list was scanned. From day 15 on, the less intensive style of encouragement characteristic of the first 9 days was resumed.

At the conclusion of the experiment, each subject was asked to read five of the lists aloud as rapidly as possible, and his total reading time for each list was timed with a stop-watch. Of the five lists read aloud, three were made up of the more common (C) words used on days 17–19 of the experiment, while the other two were from the general (G) pool. The order of presentation was CCGGC.

Results

Mean scanning speeds. Fig. 2 exhibits average time per word as a function of day of practice for the three conditions. It can be seen that subjects began by scanning fastest in *Name*, more slowly in *Animal*, and still more slowly in *N or A*, and that the three conditions converged during the course of practice as all of them improved. *N or A* was still slightly slower than the other two at the end of the experiment, but the average difference is not more than 0·01 sec/word. Individual data appear

in Table 3. During the last week of the experiment, only three subjects still showed a systematic tendency to slow down in the double condition.

The effect of restricting the lists to common words on days 17–19 was surprisingly small. Indeed, there may have been no effect at all; the slow-down on day 20 may have resulted from other factors.

The scanning rates attained in Expt. II seem to be slightly higher than those for comparable conditions of Expt. I, but the difference is small enough to be ascribable to individual variation. The substantial drop in time per word after the special encouragement on day 9 is also noteworthy.

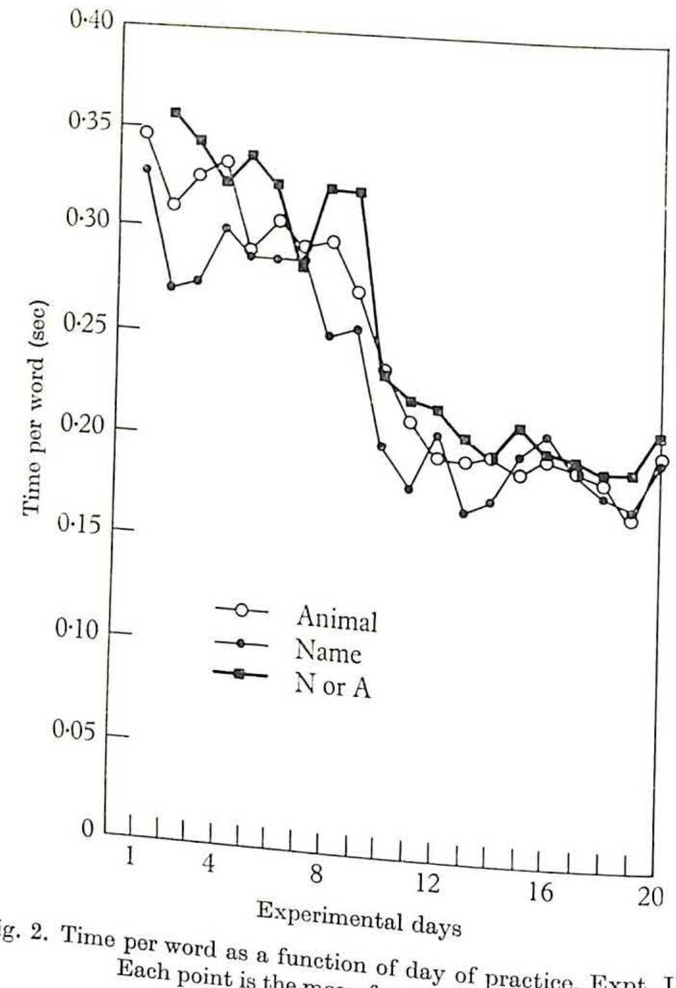


Fig. 2. Time per word as a function of day of practice, Expt. II. Each point is the mean for six subjects.

Errors. Table 4 shows that the error rates in Expt. II were slightly less than those of Expt. I, but still very substantial. Again the rate is highest in *N or A*; again there is no consistent relation between the speeds and error rates of individual subjects.

Reading aloud. Comparison of Tables 3 and 5 shows that subjects scan about twice as fast as they can read the words aloud. There is no consistent relation between the scanning rates and reading rates of individual subjects.

Table 3. *Time per word (sec) by subject and condition, Expt. II*

Condition	Subject JS	BD	EL	SS	GL	RH	Mean
Name	0·17	0·18	0·19	0·25	0·28	0·30	0·23
Animal	0·17	0·20	0·20	0·27	0·30	0·32	0·24
N or A	0·19	0·22	0·22	0·27	0·33	0·33	0·26
Mean	0·18	0·20	0·20	0·26	0·30	0·32	0·24

Table 4. *Percentage errors of omission by subject and condition, Expt. II*

Condition	Subject JS	EL	BD	SS	GL	RH	Mean
Name	12	20	15	17	14	9	15
Animal	11	18	17	16	16	9	15
N or A	14	24	16	18	22	12	18
Mean	12	20	16	17	17	10	15

Table 5. *Time per word (sec) in reading aloud, by subject and word-frequency*

Minimum frequency (per 5 million)	JS	BD	EL	SS	GL	RH
50	0·39	0·34	0·38	0·38	0·36	0·39
3	0·42	0·35	0·41	0·38	0·42	0·45

Subjective reports. No subject noticed that the lists used on days 17–19 were made up of more common words, although the difference is immediately apparent to anyone who looks for it. Five of the six reported that they usually did not become aware of the meaning of the words through which they searched.

Discussion

The results of Expt. I confirm hypothesis (1) decisively. The three conditions which require memory examination during the search process have substantially the slowest scanning rates. More time is needed to decide whether a word is a proper name than whether it is a single identifiable configuration like 'Monday'. Furthermore, the decision process is much more fallible in the former case. Within the context of a single experimental method, we have operationally defined two discrete levels of cognitive complexity.

Hypothesis (2)—that *Monday*, *States*, and *K* would give rise to equivalent scanning rates—was disconfirmed. It seems likely that the eleven possible states did not form a clearly delineated set of possible targets for our subjects. Thus, their task was more difficult than that in Neisser *et al.* (1963), where the target was any of ten letters; the possible targets in that experiment were ostensively defined, while our eleven states were not. In this connexion, it is noteworthy that one subject responded positively to 'Mass', which memory examination might classify as the name of a state since it is the common abbreviation for 'Massachusetts'. Thus we believe that subjects were occasionally examining the meanings of words in condition *states*, even though in principle they need not have done so. Apparently no such processes occurred in *Monday*.

The difficulty of condition K was quite unexpected. A preliminary experiment had shown that subjects can search for K very rapidly in lists composed of four-letter words. Indeed, the average time per word of five subjects on that occasion (averaged over trials 17–28), was only 0·09 sec, slightly less than was needed to search for K in random four-letter strings. Apparently the varying word length in the present experiment increased the effectiveness with which K was masked by neighbouring letters. In addition, subjects may have been so definitely set to look at whole words that they could not easily decompose them into constituent letters. K requires a kind of stimulus examination in which every letter must be considered, while a word-search such as *Monday* does not.

The third hypothesis, that all the searches requiring memory examination would be equally fast, was not confirmed even in Expt. II. Although the difference between *N or A* and the two single conditions became very small, and negligible for some subjects, it did not vanish entirely. The error rates also reflect the greater difficulty of the double search. However, there can be no doubt that the differences between conditions narrowed during the course of the experiment. To account for these findings, it would be necessary to develop a theory of memory structure and its use. We are not prepared to offer such a theory at present.

The tests of recognition memory carried out after Expt. I (and similar tests reported in Neisser & Stoper, 1965) indicate that subjects do not make any lasting record of the words through which they search. This suggests the existence of operationally distinguishable levels in memory examination. A subject who determines that a given word is not a name is comparing it with stored information, but not to the extent that would specify it uniquely. Hence, its occurrence as a specific individual word is not stored, nor even registered, as the introspections of our subjects indicate. This kind of incomplete memory examination also characterizes normal reading; one ordinarily becomes aware of the meaning of a passage without noting the words themselves. In some cases, one may not even know in what language the passage was written (Polanyi, 1958, p. 57). Further research will be needed to clarify the conditions under which individual words undergo more thorough cognitive analysis.

The computer tape from which our words were drawn was originally prepared by Roland Silver from Professor Lorge's punched cards, and was made available to us by Mrs Lorge. Mrs Miriam Goodman was the painstaking editor of the lists; Miss Louise Wylan ran the subjects of Expt. I; Miss Catherine Ellis conducted the preliminary experiment described in the discussion. We especially appreciate the helpfulness of the Harvard Center for Cognitive Studies, at which both the preliminary experiment and Expt. I were conducted.

REFERENCES

NEISSER, U. (1963). Decision-time without reaction-time: experiments in visual scanning. *Amer. J. Psychol.* **76**, 376–85.

NEISSER, U., NOVICK, R. & LAZAR, R. (1963). Searching for ten targets simultaneously. *Percept. mot. Skills*, **17**, 955–61.

NEISSER, U. & STOPER, A. (1965). Redirecting the search process. *Brit. J. Psychol.* **56**, 359–68.

POLANYI, M. (1958). *Personal Knowledge*. University of Chicago Press.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Teachers College, Columbia University.

(*Manuscript received* 14 *November* 1964)

Brit. J. Psychol. (1965), 56, 4, *pp.* 359–368
Printed in Great Britain

REDIRECTING THE SEARCH PROCESS*

By ULRIC NEISSER and ARNOLD STOPER
Brandeis University, Waltham, Mass.

Subjects scanned through lists of words looking for proper names. Some lists contained cue words which made it possible to skip a number of lines. One or two seconds were needed to make such a skip; hence cues for short skips were not used. Interpretation is in terms of a higher-order process which controls the scanning pattern.

In reading, one does not look carefully at every word. Successive ocular fixations, and presumably successive fixations of attention, are spaced in a way which depends partially on what has been read. In particular, one makes more fixations as the material read grows more difficult (Tinker, 1958). In other words, information that has already been assimilated determines where new information will be sought.

How quickly can such changes in the reading process be made? Perhaps only slowly: a reader who is fixating every fourth word (say) may have to read a paragraph or a page to find out that his tactics are inappropriate and that a different fixation rate might be better. At the other extreme, it may be that attention can be manœuvred on a word-by-word basis, using each word seen to direct the next movement of the eyes.

The present experiment was an attempt to determine how rapidly a visual process can be redirected. Instead of ordinary reading, we used the method of visual search (cf. Neisser, 1963; Neisser & Beller, 1965). Subjects were presented with lists of words, to be searched until a proper name was found. Some of the lists also contained cue words which indicated that a portion of the list could be skipped. We expected that these cues would be used to increase the efficiency of the search.

Method

Stimulus materials. The experimental design required two pools of English words, one of proper names and one of non-names. These were constructed from the Lorge Magazine Count in a way which has been described elsewhere (Neisser & Beller, 1965). The name pool contained 514 men's and women's first names, while the general pool contained 7901 other words. No words containing the letter K appeared in either pool, but a separate pool of 739 non-names with K, called the K-pool, was also available. All the words were between three and six letters long, and had language frequencies of at least three per five million words.

As in previous experiments, the lists to be scanned were constructed by a computer from these pools, and printed on individual pages. Standard output paper with green background lines was used. Each list was a column of fifty words drawn from the general pool, except that one of the fifty was replaced by the target word (drawn from the name pool or the K-pool, according to the condition), and in some conditions another of the fifty was replaced by a cue word, to be described below.

* This work was performed while the senior author was at the Lincoln Laboratory of the Massachusetts Institute of Technology, which operated with support from the U.S. Air Force. In addition, the research was partially supported by the National Science Foundation under grant no. G-21654 to Brandeis University. The authors also appreciate the helpfulness of the Harvard Center for Cognitive Studies, at which the experiment was conducted.

Experimental design. Two cue words, which appeared in certain conditions, could be used to redirect the search process. Whenever the word 'hop' appeared in the list, the target never appeared within the next five lines ($\frac{5}{8}$ in.). The word 'jump', on the other hand, meant that the target would not appear in the next fifteen lines ($2\frac{1}{2}$ in.). No more than one cue word appeared in a single list. These contingencies were carefully explained to subjects, who had many opportunities to confirm them during the course of the experiment. The position of the target in the list was essentially random within the restrictions imposed by the cue word, except that the target positions in certain lists were matched after day 10, as will appear below. In lists with cue words, the position of the cue was also essentially random (given the necessity of leaving room for the target) except for a similar matching procedure.

During the first ten daily sessions, there were three experimental conditions, all presented on each day. In condition *hop*, approximately two-thirds of the twenty-four lists presented per daily session contained a 'hop' cue; the others were controls. Similarly, about two-thirds of the *jump* lists contained a 'jump' cue. In *control*, twelve to fifteen lists were presented, none containing cues. Subjects knew which condition was being run. The first two lists in each condition were considered practice, and not counted as data. In *hop* and *jump* twenty-four lists were always presented, regardless of errors made by the subject. The control condition, however, was deliberately kept identical with condition *name* of Neisser & Beller (1965), so that as many as three extra trials were given to make up for errors. 'Errors' were failures to find the target in one pass through the list, or responses to non-target words.

On the eleventh day a change in procedure was introduced, for reasons given below under 'Method of analysis'. A single condition, *mixed*, was substituted for the other three. The subject scanned fifty-six lists. The first two of these (one containing 'hop' and the other 'jump') were considered practice. Of the remaining fifty-four lists, eighteen contained 'hop' and eighteen 'jump', while the other eighteen were controls. Their order was randomized, so the subject could not predict which cue, if either, he would encounter. He went through the entire set of lists regardless of the number of errors he committed.

A matching procedure ensured that the target and cue positions on twelve of the 'hop' lists were duplicated on twelve of the 'jump' lists; the same target positions were also replicated on twelve of the control lists. The actual names used as targets were randomly varied, and the order of presentation of the lists was random. Some of the other lists were also partially constrained: to ensure, for example, that the target sometimes followed 'hop' by less than the fifteen words which would be required after 'jump'.

Also beginning on the eleventh day a new condition, *K*, was introduced which was expected to produce faster scanning rates. This condition involved a separate set of twelve to fifteen lists, in which the target word was not a name, but rather contained the letter K. The procedure was otherwise like that followed in the control condition during days 1–10, and was identical with condition *K* of Neisser & Beller (1965), which was being run concurrently.

The experiment proper was run for 28 days on 5 days per week. On day 29, a special procedure was followed to determine whether the irrelevant words on such lists were committed to memory in any way. (Already on day 28 the subjects were asked about two particular words that had appeared on the two preceding lists; no subject remembered seeing either.) After the fifth list in the *mixed* condition, the experimenter read twenty-four words to the subject, who was asked which ones had appeared on the immediately preceding list (sixteen of the twenty-four actually were taken from that list, while eight were new words).

In addition, a special list was used on day 29 to determine whether the 'hop' and 'jump' skills would transfer from a search for names to a search for 'K', which previously had not been combined with cue words. Fifty-six lists were presented in a pattern analogous to condition *mixed*, but with K-words as targets.

Apparatus and procedure. The subject used a hand-held switch to turn on lights which caused the list to become visible, and immediately began to scan from the top down. The switch also started a timer, which stopped when another turn indicated that the target had been found. Details of the procedure can be found in Neisser & Beller (1965).

Subjects. One male and five female students at the Harvard Summer School were paid by the hour and run individually. One subject dropped out after day 13 for personal reasons. Her results to that point were comparable to those of the others but they are not used here in any analysis which spans the entire experiment.

The subjects knew that the experiment was an attempt to determine the maximum possible scanning speeds, and the most efficient possible use of the cue words, under the given conditions. They were repeatedly encouraged in an attempt to improve their performance.

Results

Method of analysis. Our usual method of analysing scanning data (Neisser, 1963) requires further consideration before it can be applied to such conditions as *hop*, *jump*, and *mixed*. With this method, search times are plotted against the actual positions of the targets. A computer program then finds the best-fit straight line by the method of least squares. The slope of this line is its most interesting characteristic:

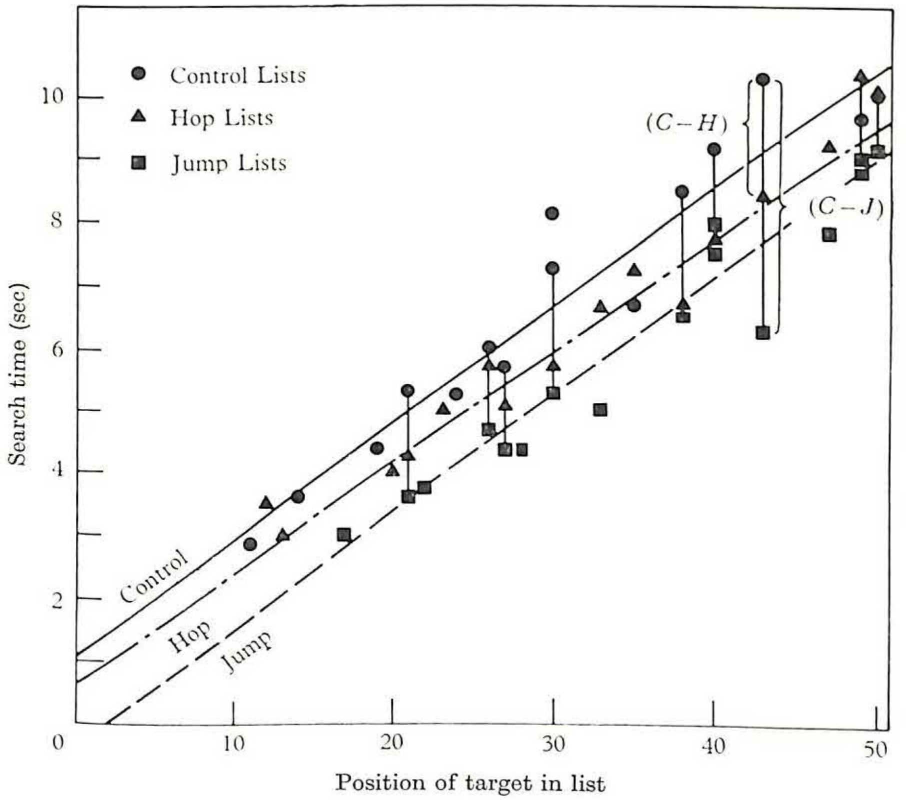


Fig. 1. Analysis of search times in condition *mixed*, day 21, subject CE. Separate lines have been fitted to points arising from the three kinds of lists. Vertical lines indicate matched points. The brackets show an example of the computations of $C - H$ and $C - J$.

it measures the increment of time needed to scan each additional word. The day-to-day variability of the slope (time per word) is usually small in practiced subjects.

It is easy to argue that the y-intercept of such a line is also meaningful. It may be thought of as a reaction-time, showing how long is needed for starting and responding, even without any words to scan. Certainly a subject who started the clock and then waited for a second before starting to scan would have an abnormally high intercept on such trials: each point would be about one second too high. By the same argument, a subject who always began with the sixteenth item, skipping the first fifteen, would produce abnormally *low* intercepts. In fact, since his line would intersect the x-axis somewhere near 15, the y-intercept would be negative.

The same considerations apply to a subject who systematically skips *any* fifteen items in the list, or any other fixed number of items, whether or not they occur at the beginning. Wherever the skip may be, provided only that it occurs before the target, its effect will be to lower the intercept of the best-fitting line. There is no *methodological* reason for expecting such skips to affect the slope. (This does not eliminate the empirical possibility that scanning may proceed at different rates in different conditions; relevant data will be considered below.) In *hop* and *jump*, some of the lists contained the cue word while others did not. It is apparent that the search-times for lists containing cue words should lie on a line with a lower intercept than those for cueless lists: hence separate lines should be fitted to these two kinds of data points. The same reasoning applies to the *mixed* condition. It is appropriate to fit three separate lines to the three kinds of points, deriving from lists containing 'jump', 'hop', or neither of these. The argument is illustrated by Fig. 1, which presents data from a practiced subject in condition *mixed*. The different intercepts and common slopes of the three lines are clearly evident.

This line of reasoning, together with Fig. 1, might suggest that the intercept is an appropriate measure for determining the differences between conditions, and the time saved as a result of a skip-cue. Unfortunately, this is not the case. The clarity of Fig. 1 is atypical. Unlike the slope, the intercept shows great day-to-day variability even in practiced subjects. It is for this reason that intercept-data have not been used in previous studies, nor can they be directly used here. The variability arises because most of the points to which the line is fitted are at a considerable x-distance from zero. Therefore, even small changes in the slope of the line are projected back into large fluctuations of the point where it intercepts the y-axis.

To obtain a better measure of the time saved by 'hop' and 'jump', condition *mixed* was introduced after day 10. Each day, the experimental design repeatedly provided the subject with three different lists that had their targets in the same position: once preceded by 'hop', once by 'jump', and once by neither. If he could save some search time by using the cues, we should expect the *absolute time* needed to find the target after 'jump' to be least, with more time needed after 'hop' and still more in the control list. The time differences between the control and jump lists, and between the control and hop lists, should indicate the extent to which the cue words were used. Each daily session produced twelve time-differences of each kind, except where this number was reduced by errors. These have been averaged to obtain mean values of $(C - J)$ and $(C - H)$ daily for each subject. It is important to see that these statistics are not scanning *rates* (time per item is about the same for all three types of lists). They correspond rather to segments of time not used for scanning. They are analogous to differences between intercepts, except that they are based directly on the observed search-times, rather than on extrapolations from them. Nine sets of matched points are indicated by vertical lines in Fig. 1, together with examples of $(C - H)$ and $(C - J)$.

Such comparisons between matched points make it possible to study the use of 'hop' and 'jump' on days 11 to 28. To make a comparable analysis for the first 10 days, the data were searched to find matched points that had arisen by chance. The random selection of target positions produced some control lists on each day which had targets in the same position as some of that day's *hop* or *jump* lists. While the

number of matched points discoverable in this way was small (an average of 3·6/subject/day) and variable, it was sufficient to provide some impression of the use of the skip-cues in the early part of the experiment.

Effect of conditions on scanning rate. Fig. 2 exhibits the daily average times per word. For the first 10 days, the three experimental conditions are plotted separately; in the *hop* and *jump* conditions, only those trials in which the cue words actually appeared were used in the determination of time per word. From day 11 on, the three separate functions are extracted from the data of the *mixed* condition. It is apparent that there are no systematic differences between conditions. This was true of the subjects considered individually also.

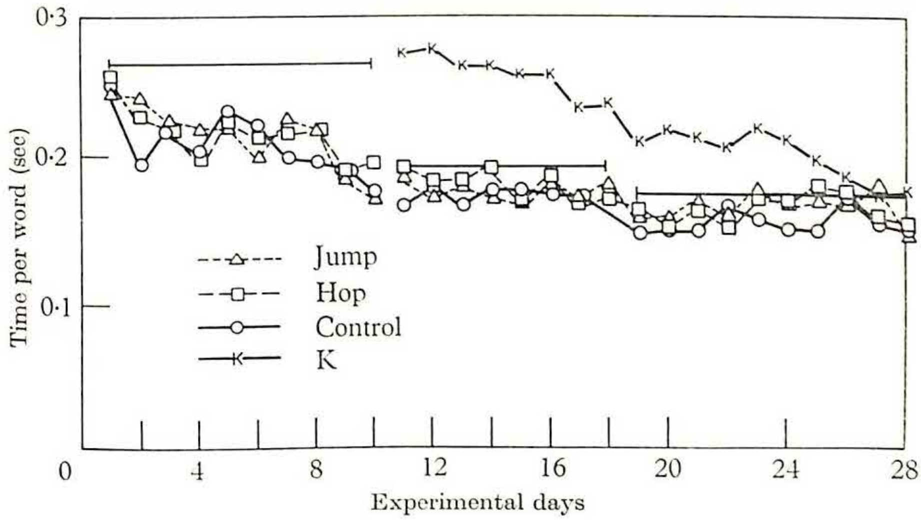


Fig. 2. Time per word as a function of day of practice. Each point represents the mean of five subjects. *Hop*, *jump*, and *control* were separate conditions on day 1–10, and subsets of a single condition thereafter. The horizontal lines represent average performances in condition *name* of Neisser & Beller (1965) over the periods indicated.

The fact that subjects scan the three kinds of lists equally fast after day 11 is hardly surprising; they could not vary their speed as a function of the cue word at least until they had scanned enough of the list to find it. The equivalence of conditions in the first 10 days is, however, more significant. If watching for a cue-word had forced the subjects to scan more slowly, times per word would have been smaller in the control condition than in the other two. Since this did not occur, we can conclude that the extra operations involved in watching for 'hop' or 'jump', and in being ready to make use of them, did not require extra time.

This conclusion is substantiated by a comparison of the scanning rates with those attained in condition *Name* (Expt. I) of Neisser & Beller (1965), which was identical with the control condition employed here. In that experiment, no skip-cues ever appeared. The three horizontal lines in Fig. 2 represent the means of six subjects in its condition *Name*, averaged over the days indicated. It is evident that our skip-ready subjects are not slower than those of the other experiment. On the contrary, they begin much faster, and only in the last week of practice do the speeds of the two experimental groups converge. We cannot be sure why the present subjects are faster initially. It is not from the advantage of 'hop' or 'jump', which does not

appear in the time per word as explained above. Individual differences may be responsible; there is some overlap between the groups.

Perhaps because they start out scanning so quickly, the subjects do not display much improvement as the experiment progresses. Time per word begins at less than

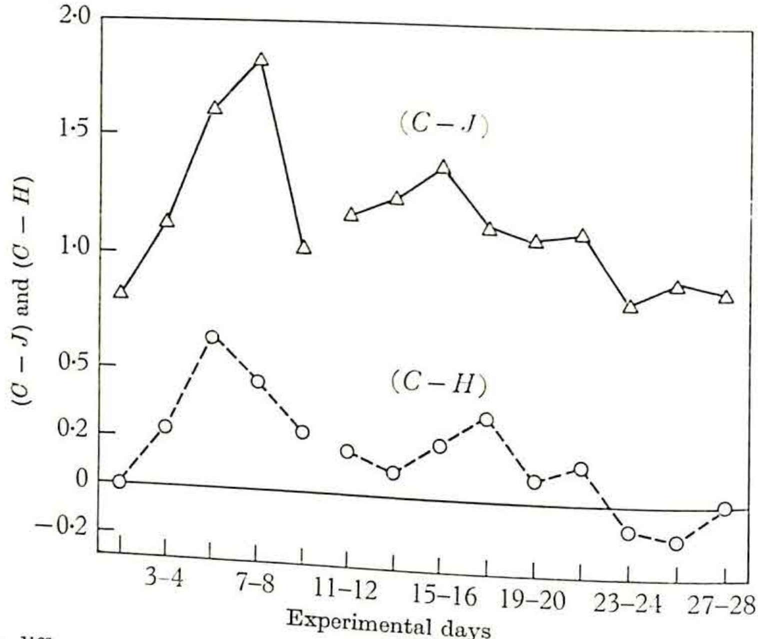


Fig. 3. Time-differences between matched points on control and 'hop' lists, and on control and 'jump' lists, as a function of day of practice. Experimental days have been paired for greater stability. The first five points on each curve are based on six subjects and about seven differences per subject; later points on five subjects and about seventeen differences for each. The curves are not independent, as both are based on differences from the same control lists.

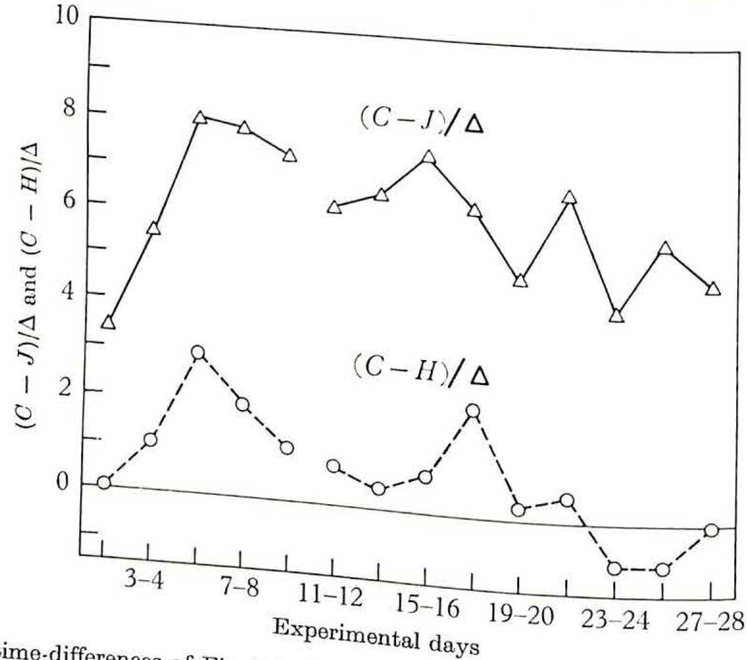


Fig. 4. The time-differences of Fig. 3 in terms of scanning rate. For each individual subject $C - J$ and $C - H$ have been divided by Δ, his time per item in the same 2-day period.

$\frac{1}{4}$ sec, and ends, with remarkable stability over the last 10 days, at about $\frac{1}{6}$ sec. One subject displayed no apparent improvement at all.

The use of 'jump'. Fig. 3 shows how the use of 'jump' and 'hop' changed during the course of the experiment. Even on the first two days, the average subject saved nearly a second of search time by using the 'jump' cue. Thereafter, the time saved by 'jumping' typically increased to a maximum before the tenth day, followed by a slow regression. At the end of the experiment, subjects were saving about as much time as they had at the beginning. This pattern, a rise in $C - J$ over the first few days followed by a decline, appeared in all subjects individually, although there was considerable variability in the magnitude of the changes. With one subject, $C - J$ dropped essentially to zero after the eighth day, but the others continued to save appreciable amounts of time with 'jump'.

Thus, there is no doubt that such a cue can be used, at the scanning speeds of this experiment. Some further analysis will help to clarify its use. For this purpose, Table 1 presents some individual data. Consider subject LR, whose scanning speed during the last ten stable days of the experiment was 0·20 sec per word (or five words per second). Over the same period, she averaged 1·79 sec less to find a target preceded by 'jump' than to find a target in the same position in a control list. This *saving* of time may be considered as the equivalent of scanning nine words. To scan through the fifteen words over which she jumped would have taken 3·00 sec; so we can say that the jump itself *cost* 1·21 sec or the equivalent of six words. Table 1 indicates that 'jumping' requires at least a second for each subject and about 1·5 sec on the average. The number of word-times saved by the jump in this concluding portion of the experiment is at most 9·0, and averages 5·3. The changes in this variable over the whole experiment is shown in Fig. 4, in which the time-differences of Fig. 3 have been divided by time per word.

The slow decrease in the time saved by 'jump' cannot be ascribed simply to an increased scanning rate, since it appears as markedly in Fig. 4 as in Fig. 3. It suggests rather that with increasing practice scanning becomes a better integrated series of actions, and thus harder to interrupt. Such an interpretation suggests that fast subjects, who presumably have better integrated scanning movements than slow ones do, should actually benefit less from 'jump'. Table 1 indicates that this is indeed the case.

The use of 'hop'. Fig. 3 and Fig. 4 show that 'hop', which indicated that only five lines could be skipped, was not well used except for a short period near the beginning of the experiment. On the average, practiced subjects save no time by using this cue. The individual data in Table 1 suggest that while the slowest subjects did make some use of 'hop', one of the fast ones was actually handicapped by it. Subject JJ took significantly more time to find the target when a 'hop' cue preceded it than in matched control lists.

Results with 'K'. This condition, in which the target was a K-word rather than a name, was introduced in the hope that it would produce a different scanning rate, at which 'hop' and 'jump' might be used in demonstrably different ways. However, as Fig. 2 indicates, asymptotic performance in this condition was not different from that in *mixed*. (It was comparable to performance in a similar condition of Neisser & Beller, 1965.) On day 29 a 'mixed-K' condition was used, in which subjects

searching for K often encountered a 'hop' or a 'jump'. No change appeared in their scanning speed. The amounts of time saved on the 'K' lists containing cue words were somewhat greater than the corresponding savings $(C - J)$ and $(C - H)$ in the regular name lists, but the difference may well have been due to chance.

Table 1. *Time per word (sec) and time saved by cues, for individual subjects*

Subject ...	LR	CE	PC	KM	JJ	Mean
Time/word (sec) (Δ)	0·20 *±0·004*	0·20 *±0·004*	0·17 *±0·004*	0·11 *±0·006*	0·14 *±0·006*	0·16
			'Jump'			
Time saved $(C - J)$	1·79 *±0·09*	1·54 *±0·09*	0·93 *±0·09*	0·65 *±0·10*	−0·19 *±0·11*	0·94
Word-times saved $(C - J)/\Delta$	9·0	7·7	5·5	5·9	−1·4	5·3
Times needed for jumping $15\Delta - (C - J)$	1·21	1·46	1·62	1·00	2·29	1·52
			'Hop'			
Time saved $(C - H)$	0·91 *±0·10*	0·27 *±0·09*	−0·07 *±0·10*	−0·05 *±0·10*	−0·28 *±0·10*	0·01
Word-times saved $(C - H)/\Delta$	1·0	1·4	−0·4	−0·5	−2·0	−0·1
Time needed for hopping $5\Delta - (C - H)$	0·81	0·73	0·92	0·60	0·98	0·81

All means averaged over days 19–28 in *mixed* condition. Italic figures are standard errors.

Table 2. *Errors by subject and type of list*

Subject ...	LR	CE	PC	KM	JJ	Total
Time/word (sec) (Δ)	0·20	0·20	0·17	0·11	0·14	—
Type of List						
Control	8	15	26	57	43	149
Hop	5	11	22	44	39	121
Jump	14	16	13	42	25	110
Total	27	42	61	143	107	380

Totals are for days 19–28 in condition *mixed*. Each subject received 18 lists of each type daily, so error percentages can be computed on a base of 180 for each internal cell, of 540 for the lower margin, of 900 for the right margin, and of 2700 (hence 14 %) overall.

A χ^2 test shows that the frequencies in the right margin depart significantly (0·05 level) from equality. The overall χ^2 shows no significant interaction between subject and type of list.

Errors. Table 2 shows wide individual differences in error rate, from 5 % for subject LR to 26 % for KM. Most of the errors, especially in the later stages of the experiment, were omissions: trials where the target was overlooked. It is interesting that appreciably fewer errors occur on lists containing 'hop' or 'jump' than on control lists. Perhaps the subject is particularly able to focus his attention on the words remaining to be searched after a 'hop' or 'jump', since the cue has indicated that the target is still to come. Table 2 also indicates that the subjects who scan fastest (and who gain the least from the cue words) make the most errors.

Recognition tests. On day 29 subjects were asked whether certain words had appeared on the list just scanned. None was able to distinguish words which had

appeared from those which had not. Some reported that no words at all looked familiar, while the accuracy of those who did make positive responses was at the chance level.

Discussion

It is evident that one can redirect a search process on the basis of cue words. However, such a change of direction takes substantial time to execute. In the present situation, the time needed appears to be between 1 and 2 sec. Since every subject was scanning at least five words/sec (300 words/min), none could make any constructive use of the 'hop' cue, which would have enabled him to skip five words. The redirection would have cost more time than it saved. (One subject, who apparently used 'hop' nevertheless, was better off in lists where it did not occur.) On the other hand, four subjects made constructive use of the 'jump' cue, which permitted a saving of fifteen words. In the time saved by using this cue, they might have scanned from five to nine additional words: the difference between these figures and fifteen leads to an estimate of 1–2 sec as the time needed to execute the jump.

This figure is quite high when compared to ordinary disjunctive reaction times, which are generally well below $\frac{1}{2}$ sec. The movement of the eye itself, as it saccades over fifteen words ($2\frac{1}{2}$ in.), accounts for only a small fraction of a second (cf. Tinker, 1958). We have eye-movement photographs of some of our subjects. For technical reasons, these photographs are hard to interpret, but they do indicate fixation rates of at least two per second; thus little time is needed to terminate one fixation and begin the next. What is taking so long?

In our opinion, these data indicate that the response (the 15-line shift of ocular fixation and of attention) is not under the direct control of the stimulus (the word as read). Eye movements in scanning and reading seem to have a pattern which is controlled by mechanisms of higher order than the immediately perceived stimulus. Lashley (1951) discusses such hierarchical control of serially ordered behaviour in a well-known paper in which he adduces such examples as the finger-movements of a pianist, too rapid to be controlled by individual sight-read notes, and syntactically controlled speech. Reaction to a cue word constitutes an interruption of an organized search pattern; time is required to bring that pattern to a halt and initiate a new one. This interpretation is confirmed by the finding that the cues take more time to use after practice than before, and are used better by slow subjects than by fast ones. Attention cannot be manœuvred on a word-by-word basis. Time is required to change a search pattern, and the change is not worth while unless it saves more time than it costs.

The data also indicate that the possibility of encountering a cue word did not hamper normal scanning. Subjects could look for both names and cue words as rapidly as for names alone, in spite of the very different responses which were contingent on finding one or the other. This confirms the finding of Neisser, Novick & Lazar (1963) that the number of things for which a subject searches can be increased without slowing the scan. The verification in this case is somewhat surprising, since a disjunctive response is involved (turn the response switch for a name, skip 15 lines for a 'jump'). However, there may be some trading relation between speed and error rate. Our subjects did make more errors in lists without cue words (in which,

therefore, they had to be alert for cue words as well as targets throughout) than in lists in which the occurrence of a cue relieved them of the necessity of looking for it any longer.

The results of the recognition tests confirm the finding of Neisser & Beller (1965) that subjects do not remember the words over which they have scanned. Although the words are 'read' to the extent necessary to determine that they aren't names, they are not 'read' so well that they can be recognized again. Scanning, and presumably reading, are processes in which the rate of input is controlled by a long-term pattern spanning more than one fixation, and in which the analysis of the input may occur with varying degrees of depth.

REFERENCES

LASHLEY, K. S. (1951). The problem of serial order in behavior. In *Cerebral Mechanisms in Behavior*, pp. 112–36. Ed. K. S. Jeffress. New York: Wiley.

NEISSER, U. (1963). Decision-time without reaction-time: experiments in visual scanning. *Amer. J. Psychol.* **76**, 376–85.

NEISSER, U. & BELLER, H. K. (1965). Searching through word lists. *Brit. J. Psychol.* **56**, 349–58.

NEISSER, U., NOVICK, R. & LAZAR, R. (1963). Searching for ten targets simultaneously. *Percept. mot. Skills*, **17**, 955–61.

TINKER, M. A. (1958). Recent studies of eye movements in reading. *Psychol. Bull.* **55**, 215–31.

(*Manuscript received* 14 *November* 1964)

Brit. J. Psychol. (1965), **56**, 4, *pp.* 369–379
Printed in Great Britain

THE EFFECTS OF REDUNDANCY AND FAMILIARITY ON TRANSLATING AND REPEATING BACK A FOREIGN AND A NATIVE LANGUAGE

By ANNE M. TREISMAN
Institute of Experimental Psychology, University of Oxford

The experiment explored the effect of sequential constraints on two speech transmission tasks: (*a*) 'shadowing' or repeating back passages in a native and in a foreign language; (*b*) simultaneous translation between a foreign and a native language. Three factors affecting performance were revealed: (1) efficiency decreased in both tasks with increased information rate in the presentation of the passages; (2) the familiarity of the language affected subjects' performance in shadowing; this effect was independent of sequential constraints; (3) translation proved more difficult than shadowing, but in this case the decrement was greater the higher the information content of the passages. This suggested (i) that sequential constraints are important in facilitating both types of speech transmission; (ii) that they are learnt concurrently with vocabulary in the acquisition of a foreign language; (iii) that translating is more difficult than shadowing not only because of decreased familiarity of either input or output, but also because of the increased decision load imposed by the more complex transformation between input and output. The roles of grammatical and semantic constraints were separated by the use of passages with words chosen at random but syntax conforming to normal rules. Performance with these passages was intermediate between that with normal prose and that with random words; the grammatical constraints seemed relatively more important than constraints of meaning in shadowing than in translating, in a native than in a foreign language, and in French than in English. The ear–voice span was measured for six subjects and was found to be greater for translating than for shadowing, but to be unaffected by the degree of sequential constraints.

Sequential constraints between words affect the efficiency with which passages varying in redundancy can be repeated back or 'shadowed' (Moray & Taylor, 1958). The percentage of words correctly repeated varied with the logarithm of the order of approximation to English. The mean information content of words in these statistical approximations to English was estimated (Treisman, 1961), and further experiments confirmed that the rate of presentation of information was the crucial factor in speech shadowing tasks. For instance, when the same passages were presented at different rates (100 words in 30 sec, 100 words in 40 sec and at rates which varied with the logarithm of the order of approximation) the percentages of words correctly repeated in each condition all fitted the same linear regression line when plotted against the rate of presentation of information (Treisman, 1961).

What happens when the speech transmission task is made more difficult or complex in other ways? Two variations on the shadowing task are of interest: (*a*) shadowing in a foreign language rather than one's native tongue, which reduces the familiarity of both stimuli and responses; (*b*) simultaneous translation between one's native and a foreign language, which may make the relation between stimuli and responses more complex, as well as reducing the familiarity of either stimuli or responses. It seemed likely that efficiency would be more markedly dependent on redundancy where the task involved a complex, arbitrary or relatively unpractised transformation from stimuli to responses, as in the second case, than where the transformation was simple but the stimuli and responses themselves were less familiar, as in the first. This is suggested by the increased decision load between input and output required

in translation: two selections need to be made, the first to identify the word or phrase heard, and the second to select an appropriate response. The shadowing task is simpler if it is assumed, as is plausible, that a single central identification of the verbal unit serves for both reception and response, so that only one decision is required.

There are some further points of interest to this problem. (i) Are sequential dependencies between words used in the same way in shadowing a foreign and a native language? A language might be learnt either as a set of word sequences or as a vocabulary acquired separately from the rules for combining its elements. In the second case one might expect the order of approximation to interact differently with performance in a foreign and in a native language. (ii) Is there any difference between translating to and from one's native language? It is usually assumed that the former is easier than the latter. If this is so, it would be interesting to see to what extent this depends on the degree of sequential constraint between words. (iii) What will be the effect of introducing purely grammatical constraints into an otherwise random sequence, so that their effect is separated from any constraints of meaning or familiar usage? This might be investigated by using passages of 'syntactical prose', constructed by choosing words at random from a novel, but with the restriction that they must have the same grammatical form as the word in the same position in a passage of normal prose taken from the same novel and serving to provide the grammatical skeleton. An experiment was designed to investigate these problems. The opportunity was taken to measure the lag between voice and ear in the different tasks, with the different degrees of sequential constraint. The eye–voice span in reading has been shown to vary with the degree of approximation to English (Lawson, 1961); would the same hold for the ear–voice span, and would this be affected by the differences between shadowing and translating?

METHOD

Passages

The 1st, 2nd, 4th, 6th and 8th order statistical approximations to English constructed by Moray & Taylor (1958) were used, together with a passage of simple English prose from a children's story and a passage of 'syntactical prose' modelled on an extract from Conrad's novel *Lord Jim*. Part of this read as follows:

> 'Up that scene the way had forgotten, maddeningly lumpily down a beard. He is perfunctorily soft with them to scatter you if he was called and held to process. I had looked at me. You were prevented for a ghost on a court.'

Each passage was 100 words long. Estimates of the mean information content per word were obtained from the frequencies with which two samples of ten words from each passage were correctly guessed by 100 subjects, presented with typescripts in which every tenth word was deleted. The frequencies were converted to probabilities and $\log_2 1/p$ was used as an estimate of the information content. For the 1st order and syntactical passages, the probabilities of those missing words which were never correctly guessed were taken from their frequencies in the Thorndike & Lorge (1944) word count. (Full details of this experiment are given by Treisman, 1965.) The values are given in Table 1, together with their conversion into the mean information rate in bits/sec for the two speeds at which the passages were presented.

A set of statistical approximations to French and a passage of syntactical French were constructed in the same way as the English passages using native French speakers or bilinguals to contribute the words (Taylor & Moray, 1960). The French prose was an extract from Colette's novel, *La Maison de Claudine*. The information content of these French passages was not estimated, as too few French subjects were available. In the experiment these values were assumed at the outset to be the same as for the corresponding English passages, except that the

prose passage was assigned an information content of 2·6 bits/word, the value found for an English extract from Conrad's *Lord Jim*, rather than the low value of 1·1 bits found for the simple 'children's English' prose passage. The results were found to be consistent with this assumption, as appears below.

Table 1. *Mean information content of English passages and rate of presentation of information in shadowing and translating tasks*

	Prose	8th order	6th order	4th order	2nd order	1st order	Syntactical prose
Mean information per word (bits)	1·1	3·3	3·6	3·9	4·8	9·2	9·4
Information rate for shadowing tasks (bits/sec)	2·75	8·25	9·00	9·75	12·00	23·00	23·50
Information rate for translating task (bits/sec)	1·84	5·51	6·01	6·51	8·02	15·36	15·70

The passages were recorded by a bilingual, female speaker on a 'Brenell' tape-recorder, at two different speeds (100 words in 40 sec and 100 words in 60 sec), and in two different orders (syntactical prose, 1st, 2nd, 4th, 6th, 8th order and prose passage, and the reverse of this). The slower speed was used for the translation task and the faster for shadowing. This difference seemed advisable, since pilot experiments showed that at the slow speed subjects made few or no errors in shadowing, while at the fast speed they correctly translated only a small proportion of words. When the number of omissions is very high, the effects of context might decrease simply because the subjects receive less of it. However, this difference in speeds meant, of course, that the absolute rate of presentation of information in bits/sec changed less rapidly over the range of passages used for translation.

Subjects

The subjects were divided into three groups of ten subjects each: (1) English speakers with a good knowledge of French; (2) French speakers with a good knowledge of English; (3) Bilinguals. The subjects in the English-speaking group were all undergraduates studying French at Oxford University. The French group were unfortunately more heterogeneous: three French people taking an Oxford Degree, three French girls spending some months in England to perfect their English, two French women who had married English members of the University, and one French 'assistante' who had been teaching in England for a year. The bilingual group consisted exclusively of people who had no preference for one language over the other; nine were undergraduates who had spent part of their lives in France and one was a French woman who had lived for many years in England.

Procedure

The subjects were given practice trials at both shadowing and translating in both languages until they appeared to be performing reasonably fluently and had ceased to show any marked improvement from trial to trial. They were then given the test passages, half the subjects in one order and half in the other and their responses were recorded. The sequence in which the four tasks—shadow English, shadow French, translate English, translate French—were given was also varied from subject to subject. The experiment took about 2 hr, 1 hr on each of two days. The responses of six of the subjects and the stimulus tape as it was played to them were recorded simultaneously on a second two-channel tape-recorder, and the ear–voice lag was measured by sampling this tape every 5 sec and noting the last word spoken on each channel. Unfortunately I thought of doing this only towards the end of the experiment, so that results are available only for the last six subjects, three of whom were bilinguals and three of whom were French.

RESULTS

The recordings of subjects' responses were analysed and the percentages of words correctly shadowed or translated was noted. An analysis of variance was then carried out, both for the whole group of subjects which allowed the variance due to native language to be estimated, and also separately for each group of subjects, in

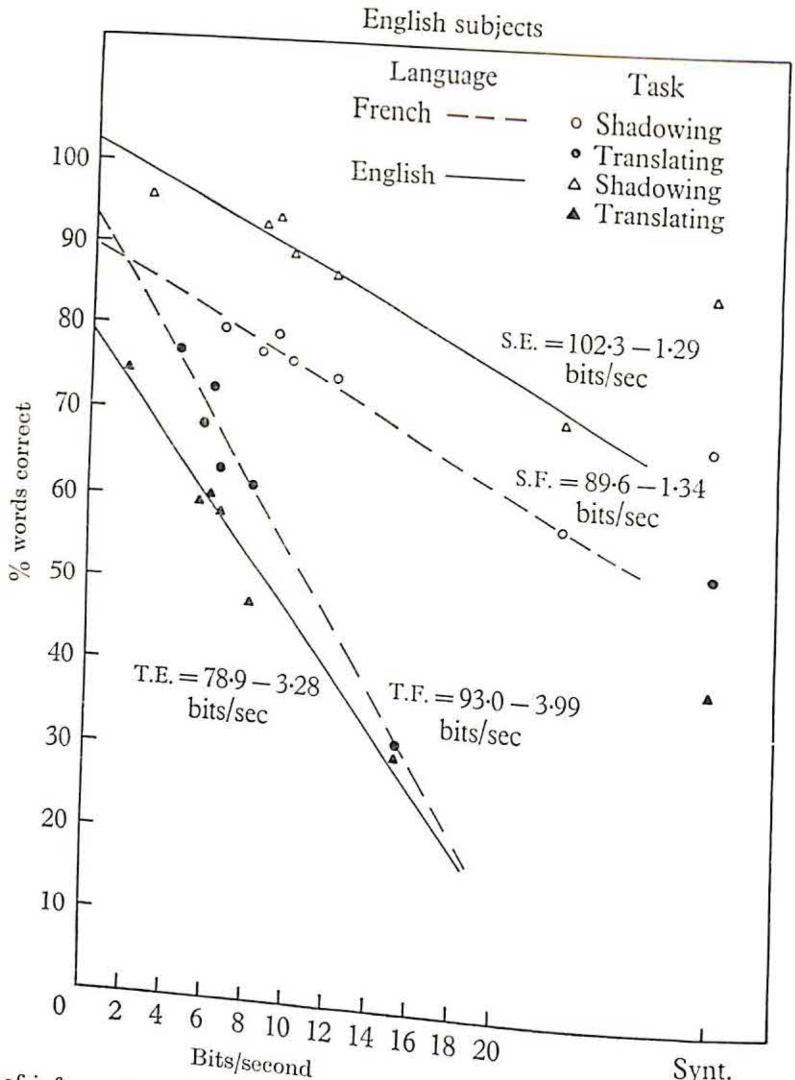


Fig. 1. Effect of information rate on percentage of correct responses in shadowing and translating for English speaking subjects with a good knowledge of French ($n = 10$).

order to compare the effects of the various factors on the results of each group. (For the analysis of variance, the percentage scores were submitted to an arc–sine transformation to ensure homogeneity of variance.) The factors analysed were subjects, order of approximation of the passages, language of the input passages, task (shadowing or translating) and native language of subjects. The passages of syntactical prose were not included in the analysis. The regression of correct responses on information rate was also calculated for the different conditions and these are shown

in Figs. 1, 2 and 3. All the regression coefficients were statistically significant and there were no significant departures from linearity. ($P < 0{\cdot}001$ was taken as an acceptable level of significance, since there was some inhomogeneity of variance in the untransformed scores.)

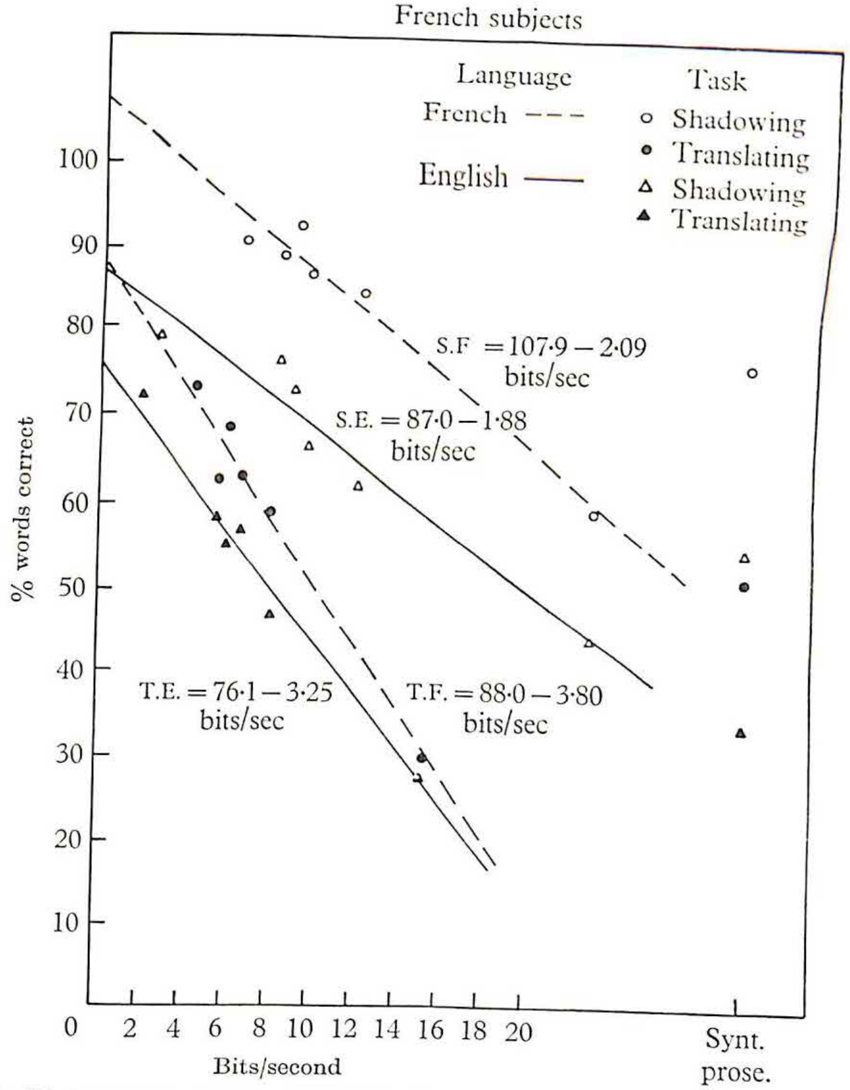


Fig. 2. Effect of information rate on percentage of correct responses in shadowing and translating for French speaking subjects with a good knowledge of English ($n = 10$).

All the main factors in the analysis of variance were significant. Efficiency decreases in both tasks as the information rate increases ($F = 510{\cdot}0$; D.F. 5, 145; $P < 0{\cdot}001$). Translating is more difficult than shadowing, even at the slower presentation rate used ($F = 2121{\cdot}5$; D.F. 1, 145; $P < 0{\cdot}001$). The bilingual group did best, followed by the English group, with the French group last ($F = 550{\cdot}4$; D.F. 2, 145; $P < 0{\cdot}001$). The language of the input passage had a significant effect on the English and French groups ($F = 35{\cdot}5$ and $270{\cdot}5$; D.F. 1, 45; $P < 0{\cdot}001$) but not on the bilinguals, confirming their equal skill in the two languages.

Perhaps the most interesting result is the interaction between order of approximation and task ($F = 11{\cdot}7$; D.F. 5, 145; $P < 0{\cdot}001$). In each group the regression of

correct responses on information rate is steeper for translating than for shadowing; for the bilinguals and the English groups, the regression coefficient is more than doubled. This contrasts markedly with the two shadowing tasks: increasing the

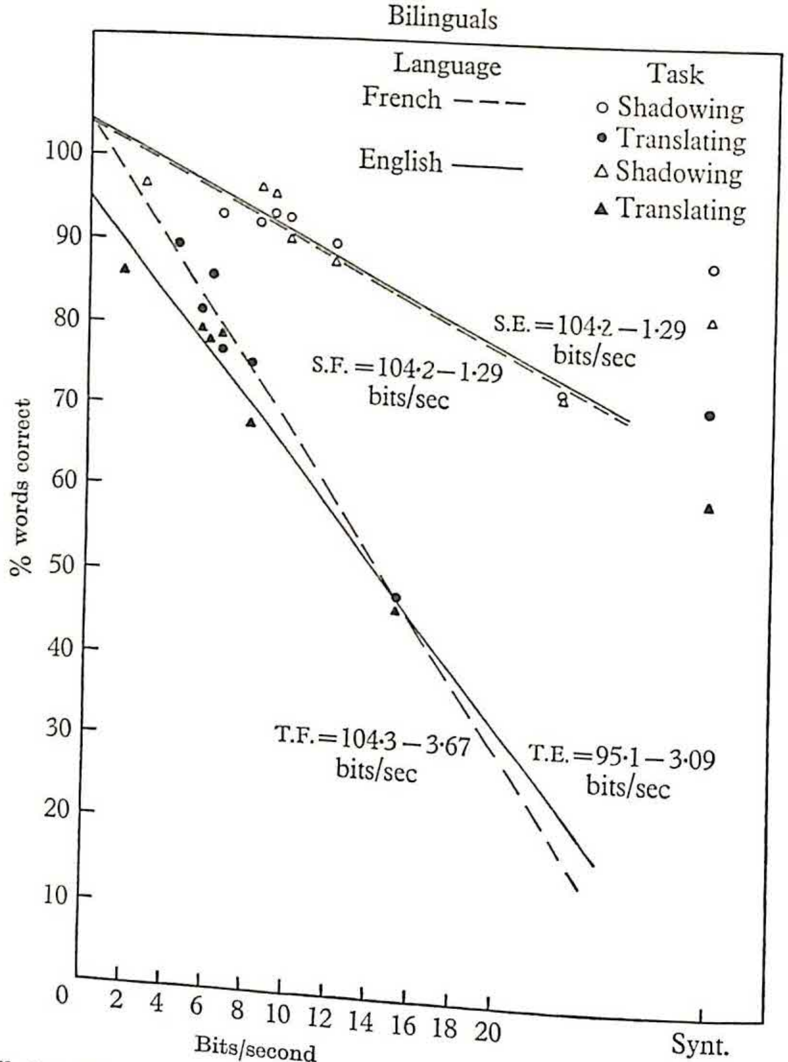


Fig. 3. Effect of information rate on percentage of correct responses for subjects bilingual in English and French ($n = 10$).

difficulty by asking subjects to shadow in a foreign rather than in their native language has no effect on the slope of the regression. It simply subtracts a constant percentage from the correct responses for all passages.

The assumption that the estimates of information content for the English passages were not too far out for the French passages is supported by the identical results obtained in shadowing each language by the bilinguals. However, all three groups show a slightly greater effect of increasing redundancy in translating French passages. The French subjects are rather more dependent on redundancy than the English or bilinguals in the shadowing task, though not in the translating.

The French and English subjects performed almost identically on the two transla-

tion tasks, both groups doing better when translating from French into English than the reverse. This was a somewhat unexpected result, in view of the widely held belief that translation is easier into than from one's native language. This was true for the English subjects, but the French subjects showed the opposite effect. The differences disappeared with the 1st order passages which were translated by both groups equally well or badly in both directions.

The results obtained with the passages of syntactical prose are shown in the graphs. The correct response rates obtained with the syntactical passages lie about half-way between those with the normal prose and those with the 1st order passage, rather closer to the 1st order for the translating task (a mean increase of 16 % compared with a decrease of 25 % from the normal prose). If they are related to the curves for the statistical approximations, the information load they present in the two tasks can be read on the abscissa. These values are given in Table 2. The overall mean estimate from these figures of the effective information conveyed in these tasks is 6·7 bits/word for the English passages and 5·8 for the French. The estimate from

Table 2. *Interpolated values for information content of syntactical passages in different tasks and languages*

Subjects	Shadow English		Shadow French		Translate English		Translate French	
	Bits/sec	Bits/word	Bits/sec	Bits/word	Bits/sec	Bits/word	Bits/sec	Bits/word
English	12·0	4·8	15·2	6·1	12·0	7·2	10·0	6·0
French	17·4	7·0	9·8	3·9	13·0	7·8	15·4	9·2
Bilinguals	16·3	6·5	11·0	4·4	11·0	6·6	8·8	5·3
Means	15·2	6·1	12·0	4·8	12·0	7·2	11·4	6·8

Table 3. *Ear–voice span in shadowing and translating*

	Span (words)				
	Shadow		Translate		Mean
	English	French	English	French	
Syntactical prose	2·6	3·3	5·3	4·8	4·0
1st order	3·2	4·1	5·0	3·9	4·1
2nd order	2·8	3·3	4·6	3·9	4·4
8th order	2·9	2·5	5·7	4·9	4·0
Prose	3·5	2·8	5·1	4·1	3·9
Mean span (words)	3·0	3·2	5·1	4·3	—
Mean span (sec)	1·2	1·3	3·1	2·6	—

subjects' guesses and the Thorndike–Lorge frequencies was considerably higher—9·4 bits/word. However, it seems that the grammatical redundancy in the syntactical passages is less helpful in translating than in shadowing, and that both shadowing and translating in a foreign language raise the effective information content of syntactical passages as compared with the subjects' native tongue.

The mean word lags or ear–voice spans in the different conditions are given in Table 3. The number of words divided by the number of words presented per second gives the span in seconds.

Analysis of variance showed no significant effect of order of approximation, but

the effects of task, language of passages and native language of subjects (French or bilingual) were all significant. There were also significant interactions between order of approximation and task and between language of passages and task ($P < 0{\cdot}001$). The ear–voice span is longer in translating English passages than French ones, though there is no difference in shadowing. The span for shadowing is shorter than the eye–voice span reported for reading by Lawson (1961), which ranged between 3·5 and 4·6 words, but similar to that reported by Chistovich, Aliakrinskii & Abul'ian (1960) for repetition of auditory speech in five out of eight subjects. Their other three subjects showed extremely small delays of about 200 msec.

Discussion

The experiment has differentiated three factors affecting the difficulty of an auditory–verbal speech transmission task. The first is the familiarity of the language to the subjects, shown in this case by the difference in shadowing a passage in a foreign and in a native language. With the foreign language there was a constant decrement over all passages, averaging 17% less correct responses. This was not a consequence of the passages used, since the bilingual group gave identical results in the two languages.

Secondly, performance decrement can also result from an increase in the information load presented: the passages with high information content and low redundancy were shadowed less efficiently than those with low information, and the relation between information rate and correct responses appeared to be linear (confirming the results of previous experiments).

Thirdly and finally, when the task is made more difficult by increasing the complexity of the transformation between stimulus and response, by asking subjects to translate rather than repeat the passages, their efficiency decreases proportionally more for the passages of high information content. The slope of the regression was doubled or trebled, suggesting that the decision load between input and output was indeed increased, as predicted, by requiring subjects both to identify the heard word and to select a different but appropriate response word. Translating must be at least a two-stage process, while shadowing may involve only one central representation shared by input and output. It would be interesting to see if the difference disappears with highly practised interpreters, where the probability of responses in one language given their translation in the other may be so high as to approach unity. In the present experiment the average difference between the intercepts on the ordinate at zero bits/sec for shadowing a native language and for translating in either direction varied with the groups of subjects: for the bilinguals, the translating task hardly affected it (an average decrease of 4%), while for the other two groups, translating reduced it by about 21% on average. This is consistent with the fact that, for the English and French subjects, the translation task not only increased the complexity of the stimulus–response relation but also decreased the familiarity of the words either for the stimuli or for the responses. It is probably no more than a coincidence that the percentages happen to sum exactly, so that the 17% reduction for unfamiliarity of words in the shadowing task plus the 4% reduction for task complexity in the bilingual group give the 21% reduction in the other two groups.

The average effects discussed above conceal some differences between the sets of passages and the groups of subjects. The French passages, with the exception of the 1st-order approximation, were consistently better translated than the English ones, by all three groups. This may be due to the structure of the language: there are more grammatical constraints in the form of agreement between adjectives and nouns, verbs and nouns, etc., in French than in English, so that passages with sequential constraints would be easier to receive and more difficult to generate in French than in English. The French group of subjects shadowed the high information passages less well than the other two groups, perhaps because they were a more heterogeneous group, whose average level of verbal fluency was probably below that of the undergraduates in the other two groups.

More surprising was the fact that their results on the translation task were identical to those of the English group, so that they found it easier to translate from their native to a foreign language than the reverse. One possible explanation is that they were living in England and much of their recent practice would have been of this sort, using English words to translate French ideas or expressions. Lambert, Havelka & Gardner (1959) report a similar finding: twenty of their forty-three subjects were faster at translating from their best reading language. They explain this by suggesting that some were 'passive' and some 'active' learners, but add that several of these twenty people were French natives at an English university, so that again the type of practice was probably asymmetrical. Further experiments, for instance on French and English groups living in France, would be needed to decide whether the higher efficiency of translation from French into English is inherent in the structure of the languages or simply an artifact of the selection of subjects. It is interesting that the difference disappears for the 1st-order passage: where translation has to be word for word and there is no sequential redundancy, it makes no difference which direction it takes. The difference between the two languages here is related to word sequences.

The presence of purely grammatical constraints and the absence of any constraints of meaning determine the results with the syntactical passages. The grammatical constraints provide assistance to subjects in shadowing and translating as compared to the 1st-order passage. This parallels the findings of Marks & Miller (1964) using a memory task, and of Miller & Isard (1963) for speech perception, both masked and unmasked. Subjects are able to use their knowledge of grammatical rules to facilitate performance on speech tasks, although the grammatical framework does not have any significant effect on the probability of the particular words in the guessing experiment which provided the estimates of word information. This suggests that at some stage between input and output, subjects must operate with a verbal framework in units other than words, even in the shadowing task, but that this will increase their efficiency only when combined with the external stimulus of particular words, and not when the verbal context alone is available and the particular word must be guessed by the subject himself. On the other hand the absence of any usual constraints of meaning or familiar usage at levels other than the grammatical considerably decreases the efficiency of performance as compared to normal prose, and even to the statistical approximations.

There is some evidence that shadowing syntactical prose in a foreign language raises the 'effective information content' compared to shadowing it in a native

language; in other words, subjects are relatively more dependent on meaning and are less able to make use of grammatical rules in an unfamiliar language than they are to use the other forms of redundancy introduced by the statistical approximations. Similarly, the syntactical constraints appear to be relatively less helpful in translating than in shadowing; the effective information is $1\frac{1}{4}$ bits less for the shadowing task. The lack of any semantic constraints seems more disruptive of efficiency in translation. This suggests that the units in translation may be elements of meaning, conceptual units, and that a translation task in which these units are reduced to single words in a semantically random sequence is more difficult than one with less grammatical constraints, but some redundancy of meaning, as in the low orders of approximations.

The main positive finding about the ear–voice span was its greater length in translating than in shadowing. The difference in size of the span again suggests that the units of speech used in translating may be longer than those adopted for the shadowing task, extending over four or five words rather than three. While straightforward repetition could theoretically be done word by word, or even phoneme by phoneme, as Chistovich *et al.* (1960) found with their subjects, this is impossible in translation, where word order differs between languages, phrases may correspond to single words, grammatical agreement may extend across several words and so on. The syntactic and semantic structure of the phrase must be known before it can be accurately translated into another language. However, this argument should not apply to the 1st-order passage, since it has in fact no inter-word dependencies. The size of the ear–voice span in this passage suggests that subjects still attempt to treat random word sequences as phrases in translating; and there were many examples of this in their responses. A typical translation into French of the 1st-order English went as follows: 'qui existe/le devoir comparé à la civilisation/de nouveau même contre les ressemblances des choses glissantes et individuelles/reconnaissant/physique/entre la supériorité/mais l'âge à la fin'—where each sequence between the strokes could be part of a sentence.

It was somewhat surprising to find that the order of approximation had no effect on the response lag or ear–voice span for the six subjects who were tested. Lawson (1961) has shown that the eye–voice span in reading varies with the degree of sequential constraint. One important difference between the two tasks is that reading is self-paced, while in the auditory–verbal tasks the rate is determined by the speed of presentation. This means that in the auditory tasks two factors may be acting in opposite directions: a high information content would tend to decrease the span if subjects try to keep the load on immediate memory constant, while on the other hand, the high information content would tend to increase the response time, so that in an externally paced task subjects would lag farther and farther behind the input. In reading they could avoid this by simply taking in the words more slowly.

I should like to thank the subjects who gave their time and skill to make these experiments possible, Mr C. Turner who analysed the tapes to determine the ear–voice span, Dr M. Treisman for his helpful comments and criticisms, and the Medical Research Council who supported this research financially.

References

Chistovich, L. A., Aliakrinskii, V. V. & Abul'ian, V. A. (1960). Time delays in speech repetition. *Probl. of Psychol.* **1**, 64–70.

Lambert, W. E., Havelka, J. & Gardner, R. L. (1959). Linguistic manifestations of bilingualism. *Amer. J. Psychol.* **72**, 77–82.

Lawson, E. A. (1961). A note on the influence of different orders of approximation to the English language upon eye–voice span. *Quart. J. exp. Psychol.* **13**, 53–5.

Marks, L. E. & Miller, G. A. (1964). The role of semantic and syntactic constraints in the memorization of English sentences. *J. verb. Learn. verb. Behav.* **3**, 1–5.

Miller, G. A. & Isard, S. (1963). Some perceptual consequences of linguistic rules. *J. verb. Learn. verb. Behav.* **2**, 217–28.

Moray, N. & Taylor, A. M. (1958). The effect of redundancy in shadowing one of two dichotic messages. *Lang. and Speech*, **1**, 102–9.

Taylor, A. M. & Moray, N. (1960). Statistical approximations to English and French. *Lang. and Speech*, **3**, 7–10.

Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Teachers College, Columbia University.

Treisman, A. M. (1961). Attention and speech. Unpublished D.Phil. thesis, Oxford University.

Treisman, A. M. (1965). Verbal responses and contextual constraints in language. *J. verb. Learn. verb. Behav.* **4**, 118–28..

(*Manuscript received* 17 *November* 1964)

Brit. J. Psychol. (1965), 56, 4, *pp.* 381–390
Printed in Great Britain

THE INFLUENCE OF LANGUAGE AND EXPERIENCE ON DISCOVERY AND USE OF LOGICAL SYMBOLS*

By HANS G. FURTH AND JAMES YOUNISS
Center for Research in Thinking and Language,
The Catholic University of America

Two non-verbal tasks of logical symbol use (SU) and of symbol discovery (SD) were devised. SU differentiated adults of higher and lower intelligence. Most pupils of chronological age (CA) 12 and older, but not younger, discovered the significance of the logical symbols in learning the paired associate task of SD. SD and SU were given to deaf subjects and to hearing subjects from a rural, lower-class milieu. These experimental subjects (CA around 18) performed similarly to controls on SU, but did not attain the success of 12-year-old controls on SD. On the basis of these results it is proposed that stimulating past experience may be required for success on certain conceptual tasks which are of the discovery type and depend on the initiative of the subject, but that more structured tasks of logical reasoning are not adversely affected by restricted experience or linguistic deficiency.

Psychological investigators have been concerned with the growth of logical thinking as it is influenced by variables which to a certain degree control its developmental manifestation. Some writers have stressed the role of language in providing the growing child with the logical categories of the environment (Brown, 1958; Luria & Yudovich, 1959; Staats & Staats, 1963); others have emphasized the role of early experience and of informal or formal training (Casler, 1961; Fowler, 1962; Hunt, 1961). In this connexion an investigation of deaf subjects' logical operations could clarify our understanding of the contribution of language in so far as the deaf in general are seriously deficient in linguistic skill and minimally exposed to the ordinary linguistic environment, at least during the early years of childhood. Language in this article refers to the living verbal language within a given society. A recent review of research with the deaf (Furth, 1964*a*) cited some studies of complex learning in which deaf performed comparably to hearing subjects. Such findings would seem to call into question a close connexion between intellective and linguistic skills. As to the retardation observed with deaf children on a number of non-verbal learning tasks, Furth suggested that restricted experience may be a more potent factor than lack of language in intellective deficit and pointed out that deaf pupils are commonly restricted in environmental stimulation, quite apart from language.

Restricted experience is a term not easily defined. It is taken here in the sense of lack of informal exposure to a stimulating environment as it is afforded within a typical educated family. In this sense children from a culturally deprived social class could be assumed to constitute a sample of children with less intellective experience than children selected from middle-class suburbia. Consequently three groups of subjects were employed in the present investigation: the deaf, as subjects

* This study was supported in part by research grant RD 1484 from the Vocational Rehabilitation Administration, U.S. Department of Health, Education and Welfare. The authors are grateful for the co-operation of the various agencies in providing subjects, particularly the Tennessee State School for the Deaf and the Board of Education, Prince George's County, Md.

not exposed to the ordinary linguistic environment and possibly also restricted in experience; socially lower-class subjects who knew language but lacked stimulating intellective experience; and controls who came predominantly from a more educated middle-class environment.

Two tasks were devised which required acquisition and utilization of certain logical principles by means of visual symbols. These symbols were similar to those used in modern symbolic logic. The tasks were non-verbal in the sense that stimuli and responses involved no connected language, and instructions were simplified so that one had reasonable assurance that deaf subjects were not penalized by their linguistic deficiency.

The *symbol use* task (SU) required comprehension of the logical symbols for affirmation, negation, conjunction and exclusive disjunction in different symbolic expressions, and application of these symbolic expressions to various objective instances. The exclusive, rather than the inclusive, disjunction was employed because it lent itself more readily to a behavioural discrimination from a conjunction. The subject's task was to indicate whether or not a given symbolic expression and an objective instance matched. It was thought that such a task would be suitable for adults and be related to a subject's intellectual ability. Furthermore, successful performance was expected to be dependent not only on the logical complexity of the symbolized expression but also on the objective context to which these logical expressions had to be applied.

The *symbol discovery* task (SD) was devised for a younger age-group and, in distinction from symbol use, was a concept attainment or discovery task. Youniss & Furth (1964) have found this task to differentiate between pupils 12–14 years old and pupils 9–11 years old. The present investigation reports the previous results in part and adds the results on this task for deaf subjects as well as for hearing subjects of lower socio-economic class.

It has been possible therefore to examine, and to compare with control data, the performance of linguistically deficient subjects (deaf subjects and lower-class hearing subjects) on two kinds of logical tasks: a task of discovery and a task of application or utilization of logical rules.

METHOD

Subjects

The older group of deaf subjects included all forty senior boys and girls at a state school for the deaf, aged 16–20 (mean CA = 18·6 yr). All forty seniors were given the symbol discovery task; for the symbol use task twenty subjects were selected from this group alternately by means of two lists which divided the seniors into those above and those below average according to school achievement. The term 'senior' refers here to deaf pupils in the age-range 16 years and above, and not to grade placement. The twenty deaf seniors who were given the symbol use task were among the group of forty deaf subjects who had had symbol discovery some time previously. Transfer effects can be ruled out, not only because of the time-interval and the difference in the requirements of the two tasks, but particularly because the deaf failed to learn the symbols in symbol discovery. The younger group of deaf subjects included all pupils at the same school aged 12–14 years (n = 27). All deaf subjects had had a severe hearing loss from early childhood which effectively hindered their learning language naturally during early childhood. None of the deaf had linguistic competence in the sense of being able to read with comprehension connected written language at the level of ordinary conversation.

The older lower-class subjects were forty-nine pupils (mean CA = 17·4 yr) from one grade of a school in a rural area. For the symbol use task ten subjects above and ten subjects below the

mean on the Lorge–Thorndike verbal scale were sampled from two class-rooms while for the symbol discovery task the remaining twenty-nine subjects from these classes were used. The younger lower-class subjects were fifty-six pupils (mean CA = 13 yr). The great majority of all these subjects came from homes where the father was an agricultural or unskilled labourer and the number of professionals or persons with college education among the parents was negligible.

The hearing control group for the symbol use task was made up of twenty-six male naval recruits (mean CA = 18·5 yr; range 17–21 yr). The high IQ group scored above average on the General Classification Test (range 49–65), and an equal number of subjects in the low IQ group scored below the average (range 28–46). The control subjects for the symbol discovery task were 309 pupils from four grades (age range 9–14 yr) of a parochial school where a majority of parents had had a college education or were professionals.

Procedure

Symbol use (*SU*). This task, consisting of six subtasks I–VI, was given individually and each trial was corrected. Two sets of cards were prepared, one set of symbol cards, the other of instance cards. On each trial one symbol card was presented together with one instance card and a subject had to give an affirmative ('yes') or negative ('no') response.

The symbolized expression on the symbol card (e.g., $B \cdot \overline{R}$) can be thought of as a logical proposition (e.g. 'There is both the colour Blue and the absence of the colour Red) or as a statement of a generalized class (e.g. 'The class of "Blue and Not-Red"'). A subject's response can be translated into a judgement of truth or falsehood concerning the logical proposition or the correctness of an instance within the logical class. As an example, a subject's behaviour could indicate the following: 'In relation to the objective event of Red and Blue, it is false to state logically the presence of Blue and the absence of Red', or, put differently: 'This particular objective event is not a possible instance of the symbolized class of "Blue and Not-Red".' Essentially then, a symbol card represented a logical expression—a mental construct which the subject had to comprehend—and an instance card presented an objective situation; a subject had to make a judgement of appropriateness in relating a conceptual and an objective event.

It should be noted that in symbolic logic the affirmative symbol (e.g. R) is here taken as $(\exists x)\cdot(Rx)$: 'It is the case that there is an x and x is red.' Similarly, the negated symbol (e.g. $\overline{R}$) is consistently interpreted as denial of the affirmation, thus $\sim(\exists x)\cdot(Rx)$: 'It is not the case that there is an x and x is red', and not as the affirmation of a class other than the negated one, $(\exists x)\cdot(\sim Rx)$: 'It is the case that there is an x and x is not red.' Consequently the instance of green is a confirmation of $\overline{R}$, not because of the presence of green but rather because of the absence of red in the instance. From this follows an apparent asymmetry when a unary (single component) symbol is matched against a binary (double component) instance card. In such a situation it is assumed (*a*) that the affirmation of a class (e.g. 'R') is instantiated when *either* component instantiates that class (e.g. 'Blue and Red'), but (*b*) that the negation of a class (e.g. '$\overline{R}$') is instantiated only when *neither* component instantiates that class (e.g. 'Blue and Yellow'). Thus the instance card, 'Red and Blue', would count as an instantiation of 'R', but it would not count as an instantiation of '$\overline{R}$'.

As our interest here was in the comprehension and application and not in the discovery of the logical symbols, subjects were told initially by words or gestures the meaning of the logical symbols, which were drawn on 3 in. × 5 in. white cards. There were four basic signs for 'Blue', 'Red', 'Circle' and 'Triangle' and they were illustrated by a 1 in. square pencilled-in with blue or red, or the black outline of a circle or a triangle, respectively. Thus it was easy to indicate, for example, that the blue square stood for 'Blue'. In a pilot study the letter 'B' was initially employed to symbolize Blue, but this caused some unnecessary verbalization, particularly with deaf subjects who considered the task to be one of vocabulary. However, the letters represent the basic signs in this report.

Six symbol cards were used for tasks I and II of the kind: B; R; $\overline{B}$; $\overline{R}$; $B \cdot R$; B/R. The dot and lines were inked black. It was pointed out that the line on $\overline{B}$ meant 'Not' (negation) and hence that $\overline{B}$ stood for 'Not Blue', that the dot between $B \cdot R$ meant 'And' (conjunction) and the slanted line meant 'Or, one or the other but not both' (exclusive disjunction). After the subject had rehearsed these instructions no furtheri nstructions or verbalizations were permitted

and task I began. Six colour instances were laid out on a table in front of the subject. These instances were made of construction paper, pasted on 3 in. × 5 in. cards and showed: Blue; Black; Red and Blue; Red; Blue and Yellow; Green and Red. The first symbol card (B) was now placed above the first instance (Blue) and the subject was required to respond with 'Yes' or 'No'. A wrong response was corrected, and a correct response was reinforced by a nod. Trial 2 followed by matching symbol B with the second instance (Black) and so forth till the sixth instance. For trials 7–12 the second symbol card (R) was used, for trials 13–18 the third ($\bar{B}$) and so on, until on the 36th and last trial of task I the symbol (B/R) was matched against instance (Green and Red).

Task II followed immediately after task I. The same material was used with reversal of procedure, each instance now being matched in turn against all the symbol cards. The examiner first laid out the previous six symbol cards and presented one instance card at a time against each of the six symbols. For trials 1–6 Blue was matched against the six symbol cards, for trials 7–12 Black, and so on.

Task III as a criterion task employed new instances and correspondingly new symbols, but introduced no new logical operations. Instead of colour, the new instances were solidly inked geometrical figures: Circle; Rectangle; Triangle and Circle; Triangle; Circle and Crescent; Square and Triangle. The symbols showed a circle (C) instead of the former (B), and Triangle (T) instead of (R). In this task a subject was given three trials on each of the following six symbolic and situational operations: (1) negation (e.g. $\bar{C}$ matched with Circle and Crescent; correct response: 'No'); (2) conjunction (e.g., $C \cdot T$ matched with Circle and Crescent; correct response: 'No'); (3) affirmation under objective conditions which presented more than the required minimum, hence related to logical simplification (e.g., C matched with Circle and Triangle; correct response: 'Yes'); and finally a symbolized exclusive disjunction (e.g., C/T) in three different objective conditions. These disjunctive conditions were: (4) 'simple' disjunction, when C/T was matched with one element alone (e.g. Circle; correct response: 'Yes'); (5) 'exclusive' disjunction in which the exclusive aspect was critical (e.g., C/T was matched with Circle and Triangle: correct response 'No'); and (6) disjunction with 'addition', when C/T was matched against one of the disjunctive forms *and* another form (e.g. Circle *and* Crescent; correct response: 'Yes'). This last operation can be seen to be related to logical addition. These six operations had been practised before in tasks I and II and were singled out in task III as possibly relevant psychological situations in which to test comprehension and application of logical symbols.

Tasks IV–VI used four new symbol cards and corresponding operations: conjunction with one negation, $C \cdot \bar{T}$ or $\bar{T} \cdot C$; conjunction with two negations, $\bar{C} \cdot \bar{T}$; negation of conjunction, $C \bar{\cdot} T$; negation of disjunction, $C \bar{/} T$; and the familiar exclusive disjunction, C/T. Tasks IV and V were presented in a manner similar to tasks I and II, respectively, i.e. for task IV instances were laid out and each symbol card, one at a time, was matched against the six instances, while for task V the six symbols were laid out and each instance, one at a time, was matched against each symbol. Task VI contained three trials each of the five logical operations of tasks IV and V, but instances and symbols changed back from form to colour.

The symbol use task as a whole was presented in two alternative forms, either starting with colour as exemplified in the above procedure, or starting with form and shifting to colour for tasks III to V and back to form on task VI. An equal number of subjects received the alternative forms. For deaf subjects on this and the following task instructions were spoken and signed at the same time.

Symbol discovery (*SD*). This task was given in group form and consisted of an attainment and a transfer task. For attainment, subjects were handed a response sheet containing numbers 1–48 and instructed to write at each presentation of coloured stimulus cards one of three possible responses written on a blackboard: $B \cdot R$; B/R and $\bar{B} \cdot \bar{R}$. Furthermore, the experimenter indicated that B stood for Blue and R for Red, but no remarks were made concerning the dots or lines that went with these letters. Six different colour combinations were shown in a predetermined series with random permutation of colours within combinations. There were sixteen cards with Blue and Red (correct response: '$B \cdot R$', i.e. conjunction); Yellow and Green (correct response: '$\bar{B} \cdot \bar{R}$', i.e. conjunction with two negations); and four of each of the following stimulus cards: Red and Yellow; Red and Green; Blue and Yellow; Blue and Green (correct response: 'B/R', i.e. disjunction). The experimenter pointed to the correct response after each trial, when all subjects had written a response.

New response sheets were used for the transfer task with three possible choices on each trial:

$C \cdot D$, C/D, and $\bar{C} \cdot \bar{D}$. Sample cards indicated to the subjects that C stood for 'Circular form', and D for 'Dark colour'. There were eighteen uncorrected trials, during which the following stimuli were presented: five cards with dark purple circular figures (correct response: '$C \cdot D$'); five with light yellow angular figures (correct response: '$\bar{C} \cdot \bar{D}$'); and four with purple angular and four with yellow circular figures (correct response: 'C/D').

Results

Symbol use

It was observed that the alternate orders of beginning with Colour or Form had no appreciable effect on subsequent performance; consequently all reported data collapse this variable. Mean errors for each logical operation for task III, and for tasks IV, V and VI combined, are presented in Table 1. While for task III these measures were based on three trials each, for tasks IV–VI they were based on 15 or 21 trials, but they have been prorated to three trials to make comparison easier. For each operation the differences between high and low IQ controls were in the expected direction. The last column of Table 1 shows mean total errors for each subject over all operations. A comparison between high and low IQ controls on these total scores gave $t = 2{\cdot}86$ (D.F. $= 24$, $P < 0{\cdot}01$) for task III and $t = 2{\cdot}30$ (D.F. $= 24$, $P < 0{\cdot}05$) for tasks IV–VI combined.

Table 1. *Mean errors of control, deaf and lower-class subjects for each logical operation*

Task III

Group	Negation	Con-junction	Simpli-fication	Disjunction: Simple	Disjunction: Exclusive	Disjunction: Addition	Total
Control							
High IQ	0·23	0·19	0·54	0·31	1·31	0·69	3·27
Low IQ	0·69	1·02	0·62	0·88	1·69	1·46	6·36
Combined	0·46	0·60	0·58	0·59	1·50	1·08	4·81
Deaf	0·25	0·55	0·35	0·40	1·80	0·85	4·20
Lower-class	0·40	0·75	0·55	0·45	1·70	0·80	4·65

Tasks IV, V and VI combined

Group	Dis-junction	Conjunction with: One Negation	Conjunction with: Two Negations	Negated: Con-junction	Negated: Dis-junction	Total
Control						
High IQ	0·53	0·50	0·42	1·06	1·41	3·91
Low IQ	1·10	0·76	0·91	1·22	1·28	5·27
Combined	0·81	0·63	0·67	1·14	1·35	4·60
Deaf	0·78	0·57	0·51	1·15	1·22	4·20
Lower-class	0·92	0·56	0·69	1·28	1·19	4·64

The error scores of the combined control group for each of the various logical operations of task III were then analysed. A rank order test for related samples was used since the error range for each single operation was small (0–3). The three disjunctive operations in general were harder than the non-disjunctive operations

(Wilcoxon $T = 38{\cdot}5$, $P < 0{\cdot}01$); moreover, within the disjunction the difficulty varied from simple disjunction and addition to exclusive disjunction (Friedman $\chi^2 = 12{\cdot}08, P < 0{\cdot}01$). The latter yielded a score comparable to chance performance. For tasks IV–VI the negated operations were harder than affirmative operations which contained one or two negated components (Wilcoxon $T = 20{\cdot}5$, $P < 0{\cdot}01$).

As can be seen in Table 1, deaf seniors performed similarly to the combined control group on the average, with their mean errors falling in the majority of instances between the high and low IQ controls. The mean performance of the lower-class subjects also closely approximated the results of the control group in single and total error scores. No differences of total mean error scores between deaf, lower-class, and controls approached significance.

Table 2. *Mean error and number of successful subjects in the Attainment Task (A) and the Transfer Task (T) of symbol discovery*

Group	All subjects n	Errors in Attainment* Conj.	Neg.	Disj.	Total	Succ. A n	Errors in Transfer* Conj.	Neg.	Disj.	Total	Succ. T n
Control											
CA 9–11 yr	169	3·93	5·96	6·21	16·10	74	2·08	2·63	5·11	9·82	8
CA 12–14 yr	140	2·97	5·68	5·88	14·53	66	0·97	0·76	2·16	3·89	42
Deaf											
CA 12–14 yr	27	4·44	6·59	6·59	17·62	15	2·80	3·30	4·90	11·00	0
CA 16–20 yr	40	3·50	4·30	5·48	13·28	26	2·50	2·50	4·50	9·50	3
Lower-class											
CA 13 yr	66	3·38	5·32	5·53	14·23	32	3·06	2·97	5·06	11·09	2
CA 17 yr	29	3·17	4·69	5·31	13·17	17	2·29	2·65	4·82	9·76	2

* Note that mean errors for attainment are based on 16 trials for each concept, and mean errors for transfer on 8 trials for disjunction and on 5 trials each for the other two concepts.

Symbol discovery

Error scores in the attainment task were based on sixteen trials for each concept, while for the transfer task they were based on eight trials for disjunction and five trials for the two other concepts. Results for the transfer task were considered only for those subjects who reached a criterion of ten consecutive correct responses in the attainment task. This explains the different total n for attainment and for transfer. The criterion for success in the transfer task was six consecutive correct choices.

Table 2 shows that a comparable proportion of older and younger controls reached the criterion for attainment (44 and 47 %). For transfer, however, inspection of the number of successful and failing subjects reveals that 64 % of the older subjects, but only 11 % of the younger subjects, performed at a criterial level which would indicate better-than-chance performance ($\chi^2 = 40{\cdot}13$, $P < 0{\cdot}001$). The deaf and lower-class subjects, aged 12–14, performed like the younger control group with minimal successes on transfer. The older deaf and older lower-class subjects also performed at the level of the younger controls and showed only about 12 % success in the transfer task. All four experimental groups, however, performed similarly to the controls in the attainment task; comparisons of the proportions of successes between the various groups gave values of χ^2 that were not significant.

To compare performance in the attainment task as a function of concept, the error scores of all successful controls ($n = 140$) were analysed by the Friedman test. The resulting $\chi^2 = 32{\cdot}3$ (D.F. 2; $P < 0{\cdot}001$) indicated that error scores varied with concept; conjunction was easier than the other two concepts.

Discussion

Symbol use

The data of the control group indicate that symbol use successfully differentiated subjects of higher and lower intelligence. The relative difficulty of exclusive disjunction compared with other operations has been reported before for tasks of discovery (Neisser & Weene, 1962) and is here confirmed in a transfer situation.

Further, performance with the same logical concept differed according to the instances with which the symbol had to be matched. This is clearly illustrated in the case of the three kinds of disjunctive operations of task III. While 'simple' disjunction required the matching of one positive component, 'addition' implied the presence of a second, irrelevant component. Finally 'exclusive' disjunction instantiated two relevant components so that the correct negative response had to be controlled by attending to the exclusive aspect of the disjunction.

Moreover, in accordance with the hypothesized interaction of concept and instance it is noted that a negated class or operation was more difficult to handle than an affirmative class which contained a negation. This may perhaps be attributed to the fact that a negated conjunction (e.g. $R \mathbin{\bar{\cdot}} B$) was exemplified by an instance (e.g. Red and Yellow) which to the perceiver directly signified the conjunction of a negated and an affirmed component ($R \cdot \bar{B}$) and only indirectly the negation of the conjunction of Red and Blue. In this connexion it is interesting also that syntactically the definition of the negated conjunction ('Not both Red and Blue') is more complex than an affirmative conjunction in which at least one component is negated ('Red and not Blue').

Symbol discovery

The attainment task consisted of six different stimuli with which three different responses had to be paired. Both older and younger controls succeeded equally well in the attainment task, but only the older controls performed better than chance in the transfer task. Apparently the twelve- to fourteen-year-old subjects, but not the nine- to eleven-year-old subjects, learned the specific relevance of the logical symbols during the attainment task and subsequently used them in the transfer task. This difference was previously related to the emergence of logical thinking after age twelve (Youniss & Furth, 1964). The older subjects were distinguished from the younger subjects by being apparently more ready to learn specific logical symbols during the attainment task, even though this task as such could be successfully mastered merely by rote learning.

Performance of the deaf and lower-class subjects

In symbol use, which has been shown to be related to intelligence, the performance of the deaf seniors was in close agreement with the normative scores. These findings support a previous study (Furth, 1964*b*) in which deaf adults were observed to

perform as well as hearing controls in classification tasks involving conjunctions and disjunctions. It seems to follow that the development of logical thinking cannot be critically dependent on the presence of language since deaf subjects who were severely deficient in language succeeded on non-verbal logical tasks as well as did hearing subjects who had had the benefit of constant exposure to language from early childhood.

It cannot be claimed that the deaf sample in this experiment consisted of unusually gifted deaf persons since the state school does not pursue a policy of selective admission and, further, care was taken to include better and poorer students in equal numbers.

These results contrast with the performance of these same deaf subjects in symbol discovery. It was unexpected that the twelve- to fourteen-year-old deaf subjects would not succeed in the transfer task of symbol discovery and it was still more surprising that deaf seniors also failed. Neither the younger nor the older deaf subjects demonstrated that they had learned the significance of the logical symbols during the attainment task of symbol discovery since they performed close to chance level in the transfer task, as did the control students of CA 9–11 yr. This marked discrepancy in performance between two tasks that presented similar concepts suggests a critical distinction between discovery which depended on the initiative of the subject, and utilization in which instructions were more precise and the task was more structured. One cannot reasonably suggest that the deaf failed to attain the logical symbols of symbol discovery because they could not master them, since they demonstrated in symbol use that they were as capable of working with these logical symbols as the controls.

In the past, the deaf have been observed to fail on some non-verbal tasks which required spontaneous shifting or restructuring and consequently they were described as being rigid (McAndrew, 1948), perceptually bound (Oléron, 1953; Oléron & Herren, 1961) and restricted (Templin, 1950) in their reasoning. Oléron (personal communication) contrasts the satisfactory performance of the deaf in structured reasoning tasks with their relative failure in tasks where more initiative is required of the subject. He and others have linked the retardation of the deaf to their obvious linguistic deficiency. Furth (1964*a*), on the other hand, has emphasized the restrictive environment and experience of the deaf and regarded language as contributing to intellectual growth in an indirect manner, in so far as language may increase the occasions for intellectual experience. Without diminishing the important role of language in social and individual development, Furth & Milgram (1965) and Hayes (1962) point out that language is not the only means of giving intellectual nourishment and it does not give it by an inherent necessity.

In this connexion the performance of the lower-class subjects in symbol use and symbol discovery is critical. These subjects came from a rural and, culturally, a rather deprived environment. In this sense of restricted experience they may be likened to the deaf. However, even though they may *use* language poorly compared with control subjects, they certainly *knew* language and had ordinary competence in their mother tongue. Thus, from a purely linguistic viewpoint, they differed radically from the deaf and only minimally, if at all, from the controls. As Tables 1 and 2 indicate, the lower-class subjects performed similarly in both tasks to the

deaf; that is, in symbol use they were as good as the controls while in symbol discovery they failed, as did the younger controls. It seems reasonable then to suggest tentatively two points. The restricted environment of the deaf and of the lower-class subjects, in contrast to the more stimulating milieu of the controls, interfered with a spontaneous deployment of logical thinking in the discovery task, while it did not obviate mature utilization of logical symbols. On the other hand, the absence of a linguistic home environment or of actual linguistic competence in the deaf did not affect negatively structured logical reasoning nor did its presence in the lower-class subjects make much difference in the case of discovery.

It may be that those subjects who went beyond rote learning in the acquisition of symbol discovery had the advantage of some internal set which facilitated spontaneous discovery of the relation between the response-symbols and the stimulus-instances. Psychologists and educators concerned with intellectual development agree on the importance of the discovery aspect of learning (Bruner, 1961) and associate discovery with factors which have been conceptualized as exploratory drive, curiosity motivation, or creative attitude (Berlyne, 1963). In so far as environmental and cultural factors have been shown to influence the manifestation of intelligent behaviour (Fraser, 1959; Tanner & Inhelder, 1957) and in the light of present findings, it is suggested that these aspects of intelligence are rather directly dependent on stimulation and on modification through specific past experiences.

In contrast, it is intriguing to consider the relative independence from effects of milieu of the more structured aspects of intelligent behaviour illustrated in symbol use. Possibly any modicum of normal experience impresses certain conclusions upon the adult so that mature forms of logical reasoning are almost universally observed, regardless of culture or formal training (Goodnow, 1962). It is in this sense that one can suggest that intelligence develops with living experience and is not dependent upon a particularly stimulating or a linguistic environment.

This proposition of the relative independence of language and thinking does not seem to be compatible with theoretical views that attribute to conventional language a uniquely important role, whether in normal development or in the education of deaf children (Lewis, 1963). Furth (1966) has enlarged on this position in its historical and theoretical context and suggests that Piaget's developmental view is consonant with the observed performance of linguistically deficient deaf subjects.

References

Berlyne, D. E. (1963). Motivational problems raised by exploratory and epistemic behavior. In *Psychology: a Study of Science*, vol. v, pp. 284–364. Ed. S. Koch. New York: McGraw-Hill.

Brown, R. (1958). *Words and Things*. Glencoe, Ill.: Free Press.

Bruner, J. S. (1961). The act of discovery. *Harvard educ. Rev.* **31**, 21–32.

Casler, L. (1961). Maternal deprivation: a critical review of the literature. *Monogr. Soc. Res. child Develop.* **26** (2, whole no. 80).

Fowler, W. (1962). Cognitive learning in infancy and early childhood. *Psychol. Bull.* **59**, 116–52.

Fraser, F. (1959). *Home Environment and the School*. University of London Press.

Furth, H. G. (1964*a*). Research with the deaf: implications for language and cognition. *Psychol. Bull.* **62**, 145–64.

Furth, H. G. (1964*b*). Conceptual performance in deaf adults. *J. abnorm. soc. Psychol.* **68**, 676–81.

FURTH, H. G. (1966). *Thinking without Language: Psychological Implications of Deafness.* New York: Free Press.

FURTH, H. G. & MILGRAM, N. A. (1965). The influence of language on classification: a theoretical model applied to normal, retarded and deaf children. *Genet. Psychol. Monogr.* (in the Press).

GOODNOW, J. J. (1962). A test of milieu effects with some of Piaget's tasks. *Psychol. Monogr.* **76** (36, whole no. 526).

HAYES, K. J. (1962). Genes, drives, and intellect. *Psychol. Rep.* **10**, 299–342.

HUNT, J. McV. (1961). *Intelligence and Experience.* New York: Ronald.

LEWIS, M. M. (1963). *Language, Thought and Personality in Infancy and Childhood.* New York: Basic Books.

LURIA, A. R. & YUDOVICH, F. I. (1959). *Speech and the Development of Mental Processes in the Child.* London: Staples.

McANDREW, H. (1948). Rigidity and isolation: a study of the deaf and the blind. *J. abnorm. soc. Psychol.* **43**, 476–94.

NEISSER, V. & WEENE, P. (1962). Hierarchies in concept attainment. *J. exp. Psychol.* **64**, 640–5.

OLÉRON, P. (1953). Conceptual thinking of the deaf. *Amer. Ann. Deaf.* **98**, 304–10.

OLÉRON, P. & HERREN, H. (1961). L'acquisition des conservations et le langage: Étude comparative sur des enfants sourds et entendants. *Enfance*, **14**, 203–19.

STAATS, A. W. & STAATS, C. K. (1963). *Complex Human Behavior.* New York: Holt, Rinehart and Winston.

TANNER, J. M. & INHELDER, B. (eds.) (1957). *Discussions on Child Development: a Consideration of the Biological, Psychological and Cultural Approaches to the Understanding of Human Development and Behaviour.* London: Tavistock.

TEMPLIN, M. C. (1950). *The Development of Reasoning in Children with Normal and Defective Hearing.* Minneapolis: University of Minnesota Press.

YOUNISS, J. & FURTH, H. G. (1964). Attainment and transfer of logical connectives. *J. educ. Psychol.* **55**, 357–61.

(*Manuscript received 2 January 1965*)

Brit. J. Psychol. (1965), 56, 4, *pp.* 391–399
Printed in Great Britain

THE RELATION BETWEEN ASSOCIATION NORMS AND WORD FREQUENCY

By W. A. MATTHEWS*
Department of Psychology, Hull University

Words of varied frequency were presented to groups of fifty subjects for a 2 min continued-association period. The mean number of associations (m values) given to the stimuli in different frequency categories failed to replicate a finding by Noble (1952*a*) whose data shows a positive relation between word frequency and m values. A possible mechanism is suggested by which these findings might be obtained and supporting analyses are presented.

Noble (1952*a*) reported an investigation in which he gave 96 disyllables to a group of 119 U.S. Air Force recruits. Each subject had to associate to each stimulus for a period of 1 min. The acceptable associations were those directly derived from the stimulus and had not to be associative chains. The mean number of associations per subject was ascertained and these ranged from 9·61 for the word 'kitchen' to 0·99 for the nonsense disyllable 'gojey'. These were labelled m values and Noble claims that they are directly related to the efficiency with which subjects learn both serial and paired associate word lists (e.g. Noble, 1952*b*; Cieutat, Stockwell & Noble, 1958).

In their attempt to find some principle underlying the learning process, Underwood & Schultz (1960) set out to develop the notion that frequency of experience is the common factor in several measures which have been used to predict learning speed.

As one example of this, they analyse in some detail the first sixty items in the table of Noble's m values. They argue that in this list a clear overall relation is found; as the number of elicited associates to a word decreases, so does its frequency of occurrence as measured by the Thorndike & Lorge (1944) G count. Certainly between the quartiles this general relation seems to hold. Nine of the fifteen words in the first quartile occur more often than fifty times per million words sampled (subsequently called 'HF words') while only five from the second quartile are HF words. The m value range sampled is from 9·61 to 5·10. In the last two quartiles, no HF words appear and the words occur less frequently than fifty times per million words sampled (subsequently called 'LF words') or fail to appear in the G count. Accordingly this could indicate that the m values are sampling an underlying frequency variable and suggests that this is the more important factor with frequency and m values being positively related. The data to be reported here casts doubt on the generality of this finding in its simple form for words at the frequent end of the m value range.

Method

Subjects

The subjects were students at a Teacher Training College (average age 20 yr). There were five groups of fifty subjects, each group consisting of twenty-five men and twenty-five women.

Material and procedure

Eighty-seven stimulus words were used. They were selected with reference to a series of experiments on short-term recall in which the associative interconnections between words

* Now at Department of Psychology, University of Southampton.

composing twelve item lists were ascertained. These experiments are reported elsewhere (Matthews, 1966).

They were presented both auditorily and visually to the subjects, who were given the following instructions: 'I am going to present a series of words to you. I will say the word, and write it on the board simultaneously. As soon as you hear and see the word, I want you to write down on the sheets which you have been given as many words as you can which this word makes you think of. Only write down those words which the word with which you have been presented brings to mind. Do not associate in a sequence from one word to another from your own responses, but after each associate has been written down always go back to the initially given word.' (An example using a stimulus word not in the set was given here to show what was and was not required.) 'To help you to refer back to the given word, always write down the stimulus as soon as you have heard and seen it in the space provided on your sheet.' (The response sheets were divided into rows and columns with the column on the left reserved for the stimulus words.) 'You will be allowed two minutes to associate to each word and then you will be told to stop. You may repeat associations if you wish to do so.' (This was clarified in response to questions but very few repetitions were in fact given.)

The responses were written on the mimeographed response sheets on which the subject had previously written the stimulus word. Each stimulus was not given to each subject. Owing to difficulties of organization five groups of subjects took part in the experiment. They were drawn from the same population and were not grouped in ways which might have influenced speed of

Table 1. *Mean number of associations given to each stimulus word in* 1 *min from a group of fifty subjects*

Stimulus	G-count word frequency	Males	Females	Total	Stimulus	G-count word frequency	Males	Females	Total
farm	AA	11·35	13·76	12·56	bats	19	7·09	9·04	8·06
kitchen	AA	9·96	11·56	10·76	poor	AA	6·91	9·12	8·01
mare	14	8·43	12·04	10·24	turnips	10	7·48	8·32	7·90
van	29	9·17	11·16	10·17	stone	AA	7·32	8·28	7·80
smoke	AA	9·04	11·08	10·06	boy	AA	7·83	7·72	7·77
log	A	9·04	11·04	10·04	eat	AA	7·76	7·68	7·72
axe	35	9·60	10·12	9·86	violin	11	7·24	8·20	7·72
orchestra	10	8·48	11·12	9·80	sentry	5	5·78	9·48	7·63
family	AA	8·48	10·96	9·72	buttercup	2	6·52	8·56	7·54
pipe	A	8·60	10·84	9·72	it	AA	7·13	7·92	7·52
cranes	9	8·80	10·16	9·48	big	AA	6·48	8·56	7·52
cold	AA	9·40	9·48	9·44	joke	32	6·84	8·04	7·46
gambling	16	9·28	9·40	9·34	show	AA	6·84	8·08	7·46
fabric	21	7·22	11·40	9·31	cool	AA	7·20	7·68	7·44
rugby	2	9·56	8·96	9·26	large	AA	6·88	7·84	7·36
rain	AA	8·56	9·64	9·10	plough	17	7·92	10·60	7·26
weave	22	7·39	10·72	9·05	Jack	A	6·72	7·80	7·26
writing	39	8·92	9·08	9·00	execution	16	6·60	7·48	7·04
trumpet	17	7·91	10·08	9·00	man	AA	5·08	8·96	7·02
life	AA	7·83	10·04	8·94	dress	AA	6·80	7·16	6·98
overture	4	7·17	10·68	8·93	wooden	46	6·92	6·88	6·90
pillow	33	7·83	10·00	8·92	legend	22	6·50	7·16	6·83
tailor	32	7·84	9·72	8·78	crates	7	6·64	6·84	6·74
island	AA	8·40	9·00	8·70	rag	28	6·08	7·28	6·68
direction	AA	7·68	9·56	8·62	article	AA	6·17	7·08	6·63
warehouse	10	7·36	9·60	8·48	suspicion	28	5·96	5·96	5·96
food	AA	7·00	9·84	8·42	giggle	5	5·20	6·68	5·94
moors	17	7·04	9·68	8·36	desires	AA	4·72	5·92	5·32
animal	AA	7·64	9·04	8·34	uncertain	21	4·76	5·88	5·32
anchor	26	7·80	8·52	8·16					

AA = occurs more often than 100 times per million words sampled; A = occur more often than 50 and less than 100 times per million words sampled; numbers, for example, 21 = occurs 21 times per million words sampled with a range from 1 to 49.

association. Group 1 was given 8 HF and 7 LF stimulus words; group 2 had 6 HF and 7 LF words; group 3 had 7 HF and 9 LF words; and group 4 had 7 HF and 8 LF words. These four groups were instructed to underline the word which had just been written at the end of 1 min, and to continue associating. The experimenter notified the group by saying 'Underline now'. Group 5 were given 27 HF and 1 LF words, in two separate sessions.

Table 2. *Mean number of associations given to each stimulus word in 2 min from a group of fifty subjects*

Stimulus	G-count word frequency	Males	Females	Total	Stimulus	G-count word frequency	Males	Females	Total
water	AA	22·80	26·20	24·50	front	AA	13·24	13·48	13·36
farm	AA	19·30	22·32	20·81	eat	AA	13·64	12·92	13·28
river	AA	16·60	20·28	18·44	moors*	17	11·60	14·80	13·20
orchestra	10	14·92	19·28	17·10	right*	AA	12·48	13·52	13·00
log*	A	15·24	18·76	17·00	bats*	19	12·30	13·64	12·97
march*	AA	15·80	17·88	16·84	boy	AA	13·52	12·40	12·96
kitchen	AA	15·72	17·44	16·58	show*	AA	13·28	12·64	12·96
light*	AA	15·16	17·44	16·30	taste	AA	11·44	14·36	12·90
horse	AA	14·32	17·92	16·12	Jack*	A	11·64	13·16	12·90
van*	29	15·30	16·68	15·99	anchor	26	11·60	13·28	12·44
smoke	AA	14·57	17·16	15·86	sentry	5	10·22	14·48	12·35
pipe*	A	14·48	17·16	15·82	stone	AA	11·96	12·72	12·34
mare	14	13·96	17·64	15·80	violin	11	11·64	13·00	12·32
cold*	AA	15·24	15·80	15·52	turnips	10	11·94	12·64	12·30
cloth	A	13·20	17·40	15·30	wooden	46	11·76	12·48	12·12
weave	22	13·13	17·44	15·29	it	AA	12·09	11·96	12·03
docks*	16	14·76	15·44	15·10	execution*	16	12·12	11·80	11·96
axe	35	15·36	14·76	15·06	point*	AA	12·36	11·48	11·92
rain	AA	14·84	15·04	14·94	buttercup	2	10·32	13·44	11·88
cranes*	9	13·68	15·84	14·76	story	AA	11·92	11·80	11·86
food	AA	12·80	16·44	14·62	legend	22	11·42	12·16	11·79
life	AA	12·83	16·40	14·61	crates	7	11·48	12·08	11·75
plough	17	12·68	16·48	14·58	needs	AA	11·40	12·04	11·72
family	AA	12·36	16·72	14·54	dress	AA	11·60	11·88	11·72
guard	AA	13·48	15·60	14·54	beast	A	10·56	12·56	11·56
house	AA	14·60	14·44	14·52	joke	32	10·08	12·92	11·50
music	AA	13·80	15·16	14·48	man	AA	8·80	13·96	11·38
fabric	21	12·13	16·40	14·26	big	AA	9·72	12·80	11·26
woman	AA	14·48	14·00	14·24	cabin	A	10·36	12·08	11·22
tailor	32	13·68	14·76	14·22	large	AA	11·20	11·20	11·20
writing	39	13·92	14·40	14·16	place	AA	10·08	12·12	11·10
gambling	16	14·44	13·76	14·10	block	A	9·60	12·32	10·96
trumpet	17	12·26	15·98	14·07	cool	AA	12·00	9·68	10·84
animal	AA	13·20	14·88	14·04	shock	AA	9·96	11·68	10·82
warehouse	10	11·92	16·04	13·98	article	AA	9·75	11·40	10·57
pillow	33	12·13	15·72	13·92	rag*	28	10·20	10·56	10·38
rugby	2	14·44	13·28	13·86	suspicion	28	9·44	9·72	9·58
poor	AA	12·04	15·48	13·76	giggle	5	7·96	10·36	9·16
island	AA	13·80	13·64	13·72	uncertain	21	8·92	9·28	9·10
night	AA	13·12	14·32	13·72	thing	AA	9·68	8·24	8·96
overture*	4	11·26	15·84	13·55	giant	A	8·40	9·16	8·78
direction	AA	12·88	14·16	13·52	desires	AA	7·20	8·84	8·02
laugh	AA	12·52	14·52	13·52	doubt	AA	6·68	7·92	7·30
field	AA	12·24	14·68	13·46					

AA = occurs more often than 100 times per million words sampled; A = occurs more often than 50 and less than 100 times per million words sampled; numbers, for example, 21 = occurs 21 times per million words sampled, with a range from 1 to 49.

RESULTS

Table 1 shows the mean number of associative responses given per subject to each stimulus during 1 min for men, women, and then combined. Thorndike & Lorge *G* count frequencies are added. Mean values are used here as these are more directly comparable with the measures reported by Noble (1952*a*). It could be argued that as associative distributions are often skewed, medians may be a more reasonable measure.

The distributions of associations to each stimulus were plotted and few instances of skewness observed. To confirm this, medians were obtained for both 1 min and 2 min scores and rank order correlations were computed between the mean and median scores. In both cases the ranks were highly correlated ($r_s >$ than $+0{\cdot}90$ for both sets of scores).

Table 2 shows the mean number of associative responses per subject to each stimulus during 2 min for men, for women, and for both combined. Thorndike & Lorge *G* count frequencies are added.

DISCUSSION

Sex differences are clearly shown in Tables 1 and 2. Women produced a greater number of associations in 54 from the 59 words used in Table 1. In Table 2, the same relation held for 71 of the 87 words listed. The decrease in the proportion of words in which the women produced a higher number of associative responses at the end of the 2 min period was not related to any one group of subjects. It suggests that although a greater number of associations may be produced by women in the earlier stages of association, any inference which suggested that women have a larger number of associations available in some hypothetical word store would be premature as this investigation lacked a necessary control measure of writing speed.

Of greater theoretical interest is the fact that Table 1 fails to show the clear relation which Underwood & Schultz (1960) claimed to have found between 1 min *m* values from Noble's list and word frequency. From the fifty-nine words listed in that table, an analysis was made of the number of HF and LF words in each quartile. Eight words from the first quartile, nine words from the second, six words from the third and nine words from the last fourteen words are LF words. This may be due to several factors which are consequential on the different subjects used. Noble used U.S.A.F. recruits. These subjects were selected Teacher Training College students in England. It is possible that the Thorndike–Lorge *G* count fails to apply to the latter population at all (rather than the inevitable failures in detail), due to different frequencies of usage of normal English words. Evidence to support or refute this suggestion is not available but the similarities in written and spoken patterns in the two countries seem to be sufficiently close to make this unlikely.

A second possibility is that the normal reading matter of students may be such that learning by a 'frequency of experience' effect may have reached some asymptote. If this were so, comparisons between HF and LF words should produce a non-ordered pattern, but the range of mean associations is surprisingly high if this were the case. Another objection may be derived from the results of a series of as yet unpublished experiments on the short-term recall of auditorily presented word-lists of twelve

items. These included both HF and LF words. Over 100 subjects took part in these experiments. They were drawn from a similar population to that which produced these norms. In eight of the groups more HF than LF words were correctly reported, although the size of the effect was not large. In the other four groups this trend was not maintained. In two of these there were no differences between the HF and LF lists; and in the two remaining, more LF words were correctly recalled than HF words. The size of this effect was also small. This indicates that word frequency is a relevant variable even though the effect is not a large one; and suggests that the Thorndike & Lorge *G* count values are reasonably valid in a verbal learning situation.

A third reason for this difference may be that these subjects were superior to those used by Noble (1952*a*) and that a time-interval of 1 min was insufficient to differentiate between the varied words. The results from the 2 min association period indicate that this is unlikely. The rank ordering in terms of *m* values for words providing the 1 min and 2 min scores correlate highly ($r_s = +0{\cdot}905$, $n = 59$, $P < 0{\cdot}001$).

A fourth possibility is that the use of different subjects for different groups of stimulus words could produce this ordering as an artifact of writing speeds or other related factors. If the ordering of the words corresponded to their grouping in presentation, this hypothesis would have been supported. An investigation of *m* values and grouping for presentation completely failed to support this hypothesis.

Excluding the subjects as a major source of variability, any search for the differences between Noble's results and these must centre on the characteristics of the stimuli. The words used here differed from those used by Noble (1952*a*) and reported by Underwood & Schultz (1960) and a comparison of the characteristics of Noble's first sixty words and those in Table 1 follows. In Noble's set, fourteen from sixty words are HF words, leaving forty-six LF words in the set. In the words in this experiment twenty-seven were HF words, and thirty-two were LF words. Fifteen of the LF words used by Noble do not appear in the *G* count. All the words used in this investigation appear in the *G* count. Thus the extreme low range used by Noble is not covered in this investigation, but otherwise the range of values used in that investigation is sampled here. However a greater variety of HF words must have been used. Differences also appear in the associative characteristics of the LF words used. Only nine of the forty-six LF words in Noble's norms were judged to have HF associates (norms are not available to demonstrate this), whereas twenty-four of the thirty-two words in this investigation were selected because they had this characteristic. Thus Noble's findings may be only applicable to a very restricted range of words, and it seems premature to generalize from a sample with these limitations to verbal processes in general.

Analyses of the data were then made to investigate some of the factors contributing to the different associative scores found in this investigation. The first analysis was of the number of grammatical usages that the exact form of the stimulus word has (e.g. noun *N*, verb *V*, adjective *A*, pronoun *P*). These classifications were made by the writer and one other person. These were then listed as triple usage words (*NAV*), double usage (*NA* or *NV* or *AV*) or single usage (*N*, *A*, *V*, *P*), and presented to twenty undergraduate students. They were asked to make rapid judgements as to the normal major use of each stimulus. Equal usage judgements were allowed. Within each usage group, the words judged to be mainly *N* or *V* or *A* were separated and the

mean number of associates to each of these classified stimuli was obtained. For a word to be placed in a group, five more individuals must have indicated that this use was more frequent than any of the others possible. However, this further breakdown revealed no differences within each usage group so that the results of this analysis will be pursued no further. The mean number of associations obtained by the variety of usage breakdown are shown in Table 3, in which the single usage words are grouped (along with a single *AV* word). This table also shows a breakdown of the four usage conditions as a function of the word frequency of the stimulus.

Table 3. *Mean number of associations as a function of usage with stimuli in word-frequency groups*

Usage	No. of words	Mean no. of assoc.	S.D.	No. of HF words	Mean no. of assocs.	S.D.	No. of LF words	Mean no. of assocs.	S.D.
3 *NAV*	24	14·86	2·09	14	15·95	1·93	10	13·37	1·56
2 *NA*	27	14·00	1·80	18	14·16	2·02	9	13·66	1·23
2 *NV*	18	11·90	2·10	13	11·08	1·78	5	12·94	2·49
1 *N* or *V* or *A* or *P* (+1 *AV*)	18	12·16	2·08	9	11·74	1·42	9	12·58	2·60

As the samples were homogeneous in variance (Bartlett test) a parametric analysis of variance was carried out on the eight conditions. The differences between the conditions were significant ($F = 8{\cdot}03$; D.F. 7; $P < 0{\cdot}001$). Comparisons were made between the individual means using the method suggested by Tukey (1953; see Ryan, 1959) and with appropriate modifications for unequal numbers in each category as suggested by Ryan (1960). It is noteworthy that this particular approach tends to maximize Type II errors. The results from these comparisons show the following significant results at $P < 0{\cdot}05$. The *NAV* HF words were significantly different from five of the seven other means (i.e. *NA* LF, *NV* HF, *NV* LF, single usage HF and LF words) but were not significantly different from *NAV* LF and *NA* HF words. No other differences attained this significance level, though several approached it. In general the significant and near-significant comparisons were between those words which included an *NA* usage and those which did not (cf. Deese, 1962, on the incidence of paradigmatic and syntagmatic associates). However, there was a clear additional effect revealed by the high scores obtained in the *NAV* HF group. Perhaps this may be attributed to frequency variables. Following the differences obtained here, a further breakdown was made in terms of words within each category which had one or more than one denotative referent (e.g. store, right, cranes). These words are starred in Table 2. Within each category there were no differences in terms of the mean number of associations produced by words which could be differentiated in this way.

The varied number of associations found between words of similar frequency could be ascribed to different numbers of situations in which each stimulus appears, or to the different number of associations that each alternative situation makes possible, or to both. If a word w appears often in one situation, it cannot appear in so many different situations as an equally frequent word ω which is not often repeated in one situation. Assuming an equal number of associations are available for each category,

this should be reflected in a wider range of different categories into which the associates of word ω could be placed, but fewer words occurring within each category group. A precise test of this hypothesis would be extremely difficult. A preliminary analysis which lacked these required controls failed to support this hypothesis. There was no control over the number of associations within a category, but an analysis was made of the individual responses to two pairs of words of similar frequency but different numbers of associations. The associations were categorized using the same criteria, and the number of categories and associations within a category counted. The words which produced the greater number of associations also contained a larger number of different categories and a greater number of words within each category. This may be because the possible associations for each category group are unequal or because different chaining methods are available as a function of the particular stimulus word.

Table 4. *An analysis of associations in common between word pairs with frequency of occurrence of the common and non-common words shown as a ratio of total occurrence of words/no. of different words*

Stimulus pairs related		No. of different words in common	Frequency with which given		Mean no. of repetitions		Mean repetitions in remaining words given as associates	
Word 1	Word 2		Word 1	Word 2	Word 1	Word 2	Word 1	Word 2
horse	mare	108	505	383	4·68	3·55	1·06	1·56
cloth	fabric	98	329	398	3·36	3·90	1·59	1·48
music	orchestra	83	353	449	4·25	5·41	1·92	1·68
cotton	weave	76	278	330	3·66	4·34	1·58	1·65
guard	sentry	69	287	314	4·16	4·55	1·80	1·66
music	overture	69	354	315	5·13	4·57	1·79	1·56
music	trumpet	62	336	252	5·42	4·06	1·73	1·96
docks	warehouse	60	294	223	4·90	3·72	2·05	1·73
field	plough	49	282	269	5·76	5·49	2·31	1·78
log	axe	44	305	300	6·93	6·82	1·75	1·96
night	pillow	43	244	257	5·67	5·98	1·81	1·83
cloth	tailor	42	197	283	4·70	6·74	1·44	2·91
field	moors	34	167	110	4·91	3·24	2·77	1·76
horse	gambling	29	135	115	4·65	3·96	2·18	1·88
field	rugby	15	88	75	5·87	5·08	2·80	2·22
Controls								
life	crates	27	82	45	3·04	1·67	1·58	1·91
island	warehouse	22	59	50	2·70	2·27	2·06	2·07
food	legend	14	29	22	2·07	1·57	2·23	1·83
march	violin	13	35	112	2·62	8·62	2·19	2·22
field	giggle	9	18	53	2·00	5·88	3·05	1·98

The last notion may be relevant to the main issue of this paper that some LF words produce as many or more associates than HF words. If the frequency norms are valid this suggests that experience in many different situations is not required to produce numerous associations. However the HF and LF words may be related to a similar context—an operational measure of this might be (*a*) a high probability of occurrence as an associate *and* (*b*) membership in the same conceptual class; if so, the associations given to the HF stimulus may be appropriate to the LF stimulus. When the LF

stimulus is presented the subject may produce associations directly derived from his experience with events which have occurred in contiguity with the LF stimulus. Associations may also be given by a controlled chaining method. If an HF word is given as an associate to the LF stimulus, associates to this HF word may then occur as associations to the LF stimulus. This could be extended to any related HF words which may occur as a response to LF associations given to the original LF stimulus. Each association obtained by this method may be produced as an overt response when it has been monitored by some mechanism as being appropriate to the LF stimulus.

Thus LF words with either HF synonyms or associated and conceptually similar HF words would be expected to produce a greater number of associations than those lacking these connexions. Examples of such pairs may be 'horse' and 'mare', and more loosely connected, 'horse' and 'gambling'. One prediction which may be made is that HF and LF words which may be paired in this way will have more words in common than randomly chosen pairs of HF and LF words of similar frequency. It is also probable that these common responses may be given more often by the group as a whole, elicited as they are by at least two stimulus events. Table 4 shows the results of an analysis made on the present data using fifteen related pairs of stimulus words and five randomly chosen control pairs. One member of each pair was an HF word. The other was an LF word. Fourteen of the fifteen pairs had more words in common than any in the control group. Additionally the associations which were produced were repeated by a higher proportion of the subjects than were the remaining associations in the response set to these stimuli. A rank order correlation was carried out on the ordering of the fifteen LF stimuli in terms of the number of associations produced and the number of different words given in common in each pair. In view of the limited sample of associated HF words available the relation was quite high ($r_s = 0{\cdot}47$, $n = 15$, $P < 0{\cdot}05$).

Additional limited support for the hypothesis may be derived from preliminary data obtained from five undergraduate subjects who produced associations verbally to a sample of these associated and conceptually related stimulus pairs and randomly selected control pairs. At both 1 min and 2 min intervals no response duplication occurred in the ten control-pair sequences which were obtained. Twenty from the twenty-five associated pair sequences produced duplicated responses. The number of words which were duplicated was not large (range 1–13) and they did not occur in similar sequences in the responses made to the two related stimulus words. It would have been surprising if this had occurred, for in the complex associative organization within subjects the preceding sequence of words must partially determine the character of the next associate (cf. Howes & Osgood, 1954).

These analyses tend to support the hypothesis put forward earlier, and to imply the inadequacy of the simple frequency of experience analysis which Underwood & Schultz (1960) suggested may underlie associative norms of this type.

This type of theoretical analysis could also be applied to explain the preliminary finding mentioned earlier, that the number of categories and number of words within a category found in associates to words of similar frequency failed to meet a prediction derived from a strict frequency analysis. The positive relation between the number of different categories and the number of associations in a category may be accounted

for by possible differences in the number of potential available associations or by the use of 'chained' associations. The greater the variety of categories into which the associations may be divided the higher is the probability that any produced association may be a stimulus for or be related to associative responses. More probably both factors are involved in these associative processes.

Thanks are due to Dr H. D. Wing, Principal of the City of Sheffield Training College, for help in providing the subjects for this investigation.

References

Cieutat, V. J., Stockwell, F. E. & Noble, C. E. (1958). The interaction of ability and amount of practice with stimulus and response meaningfulness (M.M′) in paired associate learning. *J. exp. Psychol.* **56**, 193–202.

Deese, J. (1962). Form class and the determinants of association. *J. verb. Learn. verb. Behav.* **1**, 79–94.

Howes, D. & Osgood, C. E. (1954). On the combination of associative probabilities in linguistic contexts. *Amer. J. Psychol.* **67**, 241–58.

Matthews, W. A. (1966). Continued word association and free recall. *Quart. J. exp. Psychol.* (in the Press).

Noble, C. E. (1952*a*). An analysis of meaning. *Psychol. Rev.* **59**, 421–30.

Noble, C. E. (1952*b*). The role of stimulus meaning (*m*) in serial verbal learning. *J. exp. Psychol.* **43**, 437–46.

Ryan, T. A. (1959). Multiple comparisons in psychological research. *Psychol. Bull.* **56**, 26–47.

Ryan, T. A. (1960). Significance tests for multiple comparisons of proportions, variances and other statistics. *Psychol. Bull.* **57**, 318–28.

Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Teachers College, Columbia University.

Underwood, B. J. & Schultz, R. W. (1960). *Meaningfulness and Verbal Learning*. Chicago: Lippincott.

(*Manuscript received* 3 *July* 1964)

Brit. J. Psychol. (1965), 56, 4, pp. 401–412
Printed in Great Britain

EMOTIONAL INDICATORS IN WORD ASSOCIATION

By W. P. BROWN
University of Aberdeen

Long reaction-times and other features of word-association performance have been widely accepted as indicators of emotional disturbance in subjects, but the evidence for this interpretation is scant. Indeed, there is better evidence for the view that these features indicate associative difficulty of a typically non-emotional kind. Separate analyses of the responses to thirty neutral and thirty emotional words made by 100 subjects confirmed the importance of associative difficulty in eliciting 'emotional' indicators. However, the emotionality of the stimulus also had this power to elicit indicators. Further, certain pairs of emotional indicators tended to co-occur when the stimulus was emotional to an extent which was not found with neutral stimuli. The possibility of selecting emotional stimuli for particular subjects by reference to these critical pairs of indicators is discussed, but their infrequency presents something of an obstacle.

Reviews of experiments on the perception (Brown, 1961) and recall (Rapaport, 1942) of emotional material by normal subjects disclose a widespread use of words as stimulus material and considerable reliance on associative reaction-time as a criterion of stimulus emotionality. The validity of the long reaction-time (LRT) as an emotional indicator in word association, however, has not been satisfactorily demonstrated. Indeed, Smith (1922) stated flatly that 'it is commonly recognized that reaction-time may sometimes be prolonged on account of purely "intellectual" difficulties, arising from the rarity of the stimulus-word, etc.' (p. 65). This point has been made several times since, by McClatchy & Cooper (1924), by Copeland (1932), and more recently and most cogently by Laffal (1955). Laffal's article is of especial importance because he cast equal doubt on a second emotional indicator in word association, the inability to recall one's response on a retest (reproduction failure). Previous workers such as Smith (1922) had suggested that reproduction failure was the most valid of the emotional indicators; Laffal purported to show that both LRT and reproduction failure indicated primarily non-emotional, 'intellectual' difficulties inherent in the task of finding an association to the stimulus-word. As a measure of this non-emotional associative difficulty, Laffal used the total number of different responses given to each stimulus by his group of eighty subjects. This measure correlated very highly with scores indicating the frequency of LRT and reproduction failure among responses to the same stimuli. From this he concluded that these two emotional indicators in word association 'are largely a function of the nature of the associational response hierarchy of the stimulus word' (p. 268).

From Laffal's argument, one would predict an absence of correlation between the emotional indicators and some criterion measure of stimulus emotionality which was independent of associative difficulty. A study by Levinger & Clark (1961) failed to confirm this prediction. They collected ratings of their stimulus-words on a seven-point emotionality scale, and added a GSR measure as a third emotional indicator in their experiment. A factor analysis of the intercorrelations obtained yielded two factors, identified as emotionality and associative difficulty, on both of which each of the emotional indicators (LRT, reproduction failure and GSR) had sizeable

loadings. Here, then, is evidence that emotional disturbance does show up on the emotional indicators, but, since there is ample evidence that associative difficulty also elicits the indicators, it is clearly perilous to infer emotional disturbance rather than associative difficulty from a particular instance of LRT, reproduction failure or GSR. Thus none of these indicators can be used safely to select emotional stimuli for later experimenting.

This conclusion is bound to be disquieting to experimenters wishing to use verbal materials to investigate perceptual defence, repression, etc. It is not, however, necessarily the final conclusion. A further possibility was indicated by Smith (1922) when he wrote that 'prolongation of reaction-time alone is not necessarily a complex-indicator; it is only significant if accompanied by other indicators' (p. 65). There is at least a possibility that the joint occurrence of two indicators, neither of which by itself would be good evidence of emotional disturbance, may prove to be of greater validity.

The correlational approach on which Laffal, and Levinger & Clark mainly relied is not a satisfactory tool for analysing patterns of co-occurrence among the indicators. A technique is required which will treat each response made as a separate case, rather than combining them over subjects to yield measures of stimulus-word attributes. An analysis of intrastimulus associations, not just interstimulus correlations, of pairs of indicators, is needed. An appropriate treatment for data at the intrastimulus level can be achieved using two-by-two tables (e.g. presence or absence of LRT *v.* presence or absence of reproduction failure) and χ^2 or one of its derivatives. A major investigation using this approach is the classic one by Hull & Lugoff (1921). This study, for long 'the only study on the validity of complex indicators' (Rapaport, Gill & Schafer, 1946, p. 492), suffers from an instructive deficiency in design. While Hull & Lugoff found that four of the five commonest indicators showed above chance frequencies of co-occurrence, they could not claim that this demonstrated the validity of these emotional indicators. The findings strongly suggest that the indicators share a common determinant, but cannot help to identify that common determinant as emotional disturbance or associative difficulty, or anything else, *for the want of criterion measures* of these possible determinants. It is much more difficult to provide criterion measures for an analysis at the intrastimulus level than it is when attention is confined to the grosser interstimulus differences. Indeed, it seems perverse to call for a response-specific criterion of emotionality, since it was the lack of such a measure which the word-association indicators were designed to meet. In the case of associative difficulty, however, there are brighter prospects. Laffal's measure of associative difficulty was the total number of different responses made to each stimulus. He argued, in effect, that stimuli which elicit many different responses do so because most subjects, experiencing difficulty in finding an associate at all, finally give an unusual one. But, of course, some subjects find associating to these stimuli less difficult and may give commoner responses, while a few subjects find difficulty and may give unusual responses even to an 'easy' stimulus. It follows from this that the unusualness of each *response* will give some indication of the degree of associative difficulty experienced by each subject faced with each stimulus. The most uncommon responses which occur, of course, are those given by only one subject to any particular stimulus, henceforth termed 'unique responses'. Coding each response in terms of

whether or not it is a unique response will give a rough criterion measure of associative difficulty which can be used in an intrastimulus analysis of responses. The unique response measure, to be sure, is rather inexact, since it depends upon norms which are but rough approximations to the 'true' norms. Nevertheless there is little reason to doubt that this measure will mark off a set of responses which in general have been accompanied by a relatively high degree of associative difficulty in the subjects making them. Further, it is certainly the case that Laffal's measure of associative difficulty (the total number of different responses to any stimulus) mainly depends upon the number of unique responses it elicits, and the two measures will be very highly correlated over a sample of stimuli (the correlation for the present sample of sixty words was $+0\cdot97$).

By modifying the measure of associative difficulty in this way, it becomes possible to include this variable in analyses of data at both the interstimulus and the intrastimulus levels, those of stimulus-word attributes and of separate response characteristics. Indeed, without treating the data in both these ways, the correlations reported by Laffal cannot be properly understood. He found that the number of different (and hence unique) responses, the number of LRTs and the number of reproduction failures elicited by stimulus words were all highly intercorrelated. While one might infer from other studies (Schlosberg & Heineman, 1950) that it was the unique responses which required the LRTs, this does not follow logically from the correlation: it could be that some subjects react to the associative difficulty by giving unique responses fairly promptly, while others use commoner responses but only after LRTs. Still others, perchance, give common responses quickly but forget them (and so record reproduction failures). Such interpretative uncertainty can only be reduced by examining the co-occurrence of unique response, LRT and reproduction failure among the separate responses. In this way a clearer picture of the role of associative difficulty in word-association performance may be obtained.

As noted above, there seems to be no response-specific criterion of emotional disturbance available for inclusion in an intrastimulus analysis at the response level. The emotional factor must be manipulated in a relatively gross manner. The simplest way to do this is to contrast a set of word-association data in which the emotional factor is unlikely to be operative with a comparable set of data in which emotional disturbance is likely to be present, drawing the line of demarcation between these sets along interstimulus boundaries.

The purpose of the present investigation was to take the problem of the validity of the emotional indicators a little farther than Levinger and Clark did, by adding an analysis of intrastimulus associations to the interstimulus correlation analysis, and by contrasting the results from a non-emotional word list with those from an emotional word list. The overall plan was as follows. Two lists of stimulus words were drawn up, one composed entirely of emotionally neutral words (N-words), the other entirely of words which might be emotionally disturbing (E-words). The two lists were of roughly equal associative difficulty, but within each list a wide range of associative difficulty levels was represented. The two lists were combined for the purpose of testing subjects, but the responses to each list were analysed separately. Analyses of the N-word responses were designed to disclose what relations among the various indicators occurred in the absence of the emotional determinant, including

those which could be attributed to the associative difficulty factor. Analyses of the E-word responses were intended to show what relations appeared when both determining factors were at work. A comparison of the results from the N-word and E-word analyses was undertaken to indicate what relations, if any, were uniquely dependent on the influence of the emotional factor.

Method

Subjects

The subjects were 100 undergraduate students, 70 women and 30 men, enrolled in the first-year course in psychology at the University of Aberdeen during session 1957–58.

Procedure

A 100-word stimulus list was used. Stimuli were presented visually in a brightly lit Dodge-type tachistoscope, with exposure durations of about 0·7 sec (the longest available with this apparatus). The subjects were instructed to respond as quickly as possible, saying the first word that occurred to them on seeing the stimulus. Responses and reaction-times (measured to the nearest 0·1 sec) were recorded. Immediately following the end of the list, a retest was given, subjects being instructed to recall the response given on first presentation. To save time, the experimenter used oral presentation of the stimuli on the retest, and merely noted whether recall was correct, incorrect, or absent.

Stimulus words

All stimulus words were of six letters, and all had a frequency of at least 10/million printed words in the general count (Thorndike & Lorge, 1944). Of the original 100 words, 60 had been chosen as E-words and 40 as N-words on *a priori* grounds. For the present analyses only 60 words were retained, 30 of each type. No words from either the earliest or the last sections of the list were retained, since words in these favourable serial positions tend to attract unduly few LRTs and reproduction failures respectively (Hubbard, 1924). The main principle of selection, however, was the requirement that the two groups of E-words and N-words should be of roughly equivalent associative difficulty (see Table 1).

While the original classification into E-words and N-words was an *a priori* one, it has subsequently received some empirical confirmation. A group of fifteen women students at Bedford College rated all 100 words on a seven-point emotional–unemotional scale. For the 60 words retained for analysis, the biserial r between mean ratings and the original classification was $+0{\cdot}89$.

The N-words were: detail, patent, person, action, theory, agency, lawyer, notice, manner, obtain, column, estate, fringe, museum, needle, shadow, employ, vision, centre, bright, branch, comedy, sphere, errand, planet, autumn, warmth, foster, garage, rustle.

The E-words were: forbid, grudge, modest, normal, excuse, attack, offend, menace, secret, vulgar, gossip, coward, caught, secure, decent, caress, choose, strict, stupid, wicked, gamble, health, doctor, vanity, guilty, afraid, hatred, sister, wealth, female.

Both lists are in order of decreasing associative difficulty (number of unique responses), except that ties are not indicated.

Indicators scored

Initially ten indicators were counted, but four of these (failure to respond, clang associations, repeating the stimulus before responding, and responding with a grammatical variant of the stimulus) occurred too infrequently to permit statistical analysis. The remaining six indicators were as follows.

Unique Response. Any response made by only one of the 100 subjects to a particular stimulus was scored 'unique response'. This indicator, as explained above, served as the criterion measure of associative difficulty. The other five indicators are all allegedly emotional indicators.

Long Reaction Time. Following Hull & Lugoff (1921) and Laffal (1955), a reaction-time of

2·6 sec or longer was scored LRT. Despite the visual mode of presentation, this cut-off point classified as LRTs almost exactly the same proportion of responses as in Laffal's study.

Perseveration. Two indicators of perseveration were used. These were Response repetition, which was scored when a subject responded with a word already used as a response to a previous stimulus, and Stimulus repetition, when a subject used a previous stimulus from the list as a response. The rationale underlying the proposal that perseveration is an emotional indicator assumes that a perseverative response is not a true associate of the stimulus. When different stimulus-words which are closely connected with one another appear in the same stimulus list, true associations may occur which meet the above definitions of R-repetition and S-repetition; in an attempt to exclude such cases, no response given by more than five of the 100 subjects was regarded as an instance of either R-repetition or S-repetition. It should be noted that these two indicators are independent and could, and did, co-occur.

Reproduction failures. Two types of reproduction failure were distinguished. These were forgetting, when the subject was unable to recall his previous response at all, and misremembering, when the subject recalled the previous response incorrectly. Forgetting and misremembering, of course, are mutually incompatible, and cannot co-occur. In scoring for reproduction failures maximum credit was given to subjects, so that if a subject made a correct recall but withdrew it in favour of an incorrect recall or an avowal of forgetting, correct recall was scored nevertheless. (The two types of reproduction failure may be equivalent, but it seemed inadvisable to assume this in advance. Previous practice has varied; Laffal apparently counted both forgetting and misremembering as failures, Levinger and Clark only counted forgetting.)

Methods of analysis

Two tables of 3000 cells (30 stimulus words by 100 subjects) were drawn up, one each for N-words and E-words. Code numbers for the six indicators were entered in appropriate cells. Row totals, showing the frequency of each indicator in response to each stimulus, were intercorrelated for each pair of indicators. In addition, the 3000 cells of the original tabulation were scrutinized, and two-by-two tables for each pair of indicators were drawn up to show the observed frequencies of joint occurrence, separate occurrence, and non-occurrence. The significance of these patterns of co-occurrence was assessed by the χ^2 technique.

Finding expected frequencies for these tables presented some difficulties. The calculation of independence values assumes that the chance probability of each indicator is the same in each of the 3000 cells. In fact, the probability was demonstrably higher in some rows of cells (i.e. to some stimuli) than in others. When both indicators involved have higher, and lower, chance probabilities in the same rows as one another (i.e. when there is significant interstimulus correlation), the normal independence values will underestimate the true f_e's of co-occurrence over the entire 3000 cells. (The study by Hull & Lugoff (1921) is open to criticism in this respect.) An attempt to avoid the danger of accepting insignificant associations as significant was made by calculating the f_e's separately for each row of the original tabulation, and summing the thirty values so obtained to arrive at f_e's with which to compare the f_o's counted up over the entire 3000 cells. In information-theory terms (Attneave, 1959), an attempt was made to assess the significance of $\hat{T}_{\text{stimuli}}(x;y)$ rather than $\hat{T}(x;y)$.

Results

Frequency of indicators

Table 1 shows how the six indicators were distributed among the stimulus words of the two lists, and also gives the mean frequencies of the indicators among responses to N-words and E-words. Since the lists were chosen to be comparable in associative difficulty, they did not differ significantly in the mean frequency of unique responses (or in the variance of this measure). All five emotional indicators, however, were elicited more frequently by E-words than by N-words. This agrees with the findings of Levinger & Clark (1961) for LRT and forgetting. If the reliability of the difference between E-words and N-words is taken as a criterion, forgetting emerges as the strongest, LRT as the weakest indicator of emotion.

Table 1. *Distribution and mean frequency (percentage incidence) of six indicators among responses to neutral and emotional words*

Indicator	Word list	Frequency of indicators among 100 responses: 0–5	6–11	12–17	18–23	24–29	30+	Mean	t
Unique R	N	4	9	3	3	6	5	17·8	1·39
Unique R	E	3	1	7	6	7	6	21·7	
LRT	N	9	8	2	7	1	3	12·2	2·38*
LRT	E	2	7	5	8	5	3	18·4	
R-repetition	N	19	11	—	—	—	—	4·2	3·12**
R-repetition	E	11	15	4	—	—	—	7·2	
S-repetition	N	27	3	—	—	—	—	1·5	2·90**
S-repetition	E	22	8	—	—	—	—	3·5	
Forgetting	N	16	8	3	3	—	—	6·7	4·07**
Forgetting	E	6	5	11	6	2	—	13·7	
Misremembering	N	22	8	—	—	—	—	4·0	2·60*
Misremembering	E	15	10	5	—	—	—	6·3	

* Difference between means significant at 0·05 level.
** Difference between means significant at 0·01 level.

Table 2. *Measures of interstimulus correlation and intrastimulus association between pairs of indicators*

Interstimulus correlations[1]: N words	E words	Pairs of indicators	Intrastimulus associations[2]: N words	E words
0·76**	0·76**	Unique R × LRT	41·18**	99·43**
0·47*	0·67**	Unique R × R-repetition	23·74**	48·48**
0·75**	0·37	Unique R × S-repetition	15·20**	7·67*
0·59**	0·62**	Unique R × forgetting	30·40**	69·03**
0·50*	0·09	Unique R × misremembering	16·21**	22·68**
0·55*	0·49*	LRT × R-repetition	0·50	0·04
0·72**	0·03	LRT × S-repetition	0·01	1·63
0·61**	0·68**	LRT × forgetting	5·51	31·81**
0·48*	0·11	LRT × misremembering	0·10	3·08
0·80**	0·33	R-repetition × S-repetition	1·92	15·29**
0·50*	0·54*	R-repetition × forgetting	5·76	12·53**
0·37	0·10	R-repetition × misremembering	4·60	13·70**
0·59**	0·11	S-repetition × forgetting	0·89	4·44
0·38	0·22	S-repetition × misremembering	11·05**	0·03
0·65**	0·33	Forgetting × misremembering	—	—

[1] When S-repetition is a variable, entries are biserial correlations; otherwise, entries are product-moment correlations, with 28 D.F.
[2] Entries are values of χ^2, with 1 D.F.

* $P < 0{\cdot}01$. ** $P < 0{\cdot}001$.

Interstimulus correlations between indicators

The correlations between each pair of indicators are shown on the left of Table 2, for N-words and E-words separately. Since fifteen correlations were derived from each set of data, little reliance can be attached to the 0·05 level of significance, and only correlations significant at the 0·01 level are to be considered substantive.

All but two of the fifteen correlations relating to N-words reached this level of significance. Among N-words, it seems, stimuli which attract any one indicator are

likely to attract all the others. The only exception to this generalization is that stimuli which attracted many perseverative responses (R-repetition and S-repetition) did not suffer markedly from misremembering of responses on the retest, although they did suffer from forgetting of them. Despite these two non-significant but not negligible correlations, the overall impression is clearly one of the total intercorrelation of all variables. Associative difficulty (unique responses) was as closely related to the emotional indicators as they were to one another.

The correlations relating to E-words present a somewhat different picture. Only six of the fifteen were significant. Neither S-repetition nor misremembering correlated significantly with any other indicator; all the correlations were significant among the four remaining variables (unique response, LRT, R-repetition, forgetting). It is difficult to interpret this pattern of intercorrelations as it stands; that is, without reference to the analysis of intrastimulus associations. It seems clear, however, that the introduction of the emotional factor had an effect on the matrix. None of the indicators functioned exactly as it did in the N-word matrix. Unfortunately, from the correlations alone, it is impossible to tell which of the variables had changed its meaning, although it is apparent that some, if not all, of them must have changed.

Intrastimulus associations between indicators

Information about the intrastimulus patterns of co-occurrence of pairs of indicators is given on the right of Table 2, for N-words and E-words separately. The entries are values of χ^2 derived from analysing the two-by-two tables, and again the criterion of significance used is the 0·01 level. All significant values refer to above-chance frequencies of co-occurrence, although in some of the insignificant cases the observed frequencies were marginally below chance.

Six of the fourteen χ^2 values relating to N-words were highly significant. Unique responses were associated with all five emotional indicators. In other words, unique responses typically have LRTs and are forgotten or misremembered on the retest, while perseverative responses (both R-repetitions and S-repetitions) are likely to be unique responses. This finding fully accords with Laffal's (1955) claim that the so-called emotional indicators are highly sensitive to associative difficulty of a non-emotional kind. The only other significant χ^2 relating to N-words shows that S-repetitions tended to be misremembered on the retest.

There are nine significant χ^2 values in the E-word analysis. Five of these relate to the associations between unique responses and the emotional indicators. Associative difficulty, then, gave rise to emotional indicators in responses to E-words and N-words alike. But there remain four significant χ^2 values for the E-word responses which were not significant for N-word responses. Responses which had LRTs were likely to be forgotten on the retest; R-repetitions were likely to be forgotten or misremembered; and R-repetitions were likely to be S-repetitions too (i.e. once a stimulus had been given as a response, it was likely to be used as a response again). None of these statements holds for N-word responses; they are all true of E-word responses. That is, these statements are true only of the set of data in which the emotional factor was believed to be operative. It may be, then, that these four pairs of indicators are indicative of emotional disturbance in a way in which no single indicator is.

Comparison of the two analyses

The two analyses, of interstimulus correlations and intrastimulus associations, yielded clearly different patterns of relation among the indicators. Among N-words, a few significant associations gave rise to many significant correlations; among E-words, a larger number of associations supported a smaller number of correlations. It is necessary to consider the different types of outcome in more detail.

Both correlation and association significant. In these cases, it can be presumed that the significant association between the paired indicators accounts for the significant correlation. Thus, for example, the explanation of the fact that stimuli which elicited most unique responses also elicited most LRTs is that many of the unique responses themselves required LRTs. For N-words, this simple pattern of relationship held between unique response and each of the five emotional indicators. For E-words, the pattern occurred five times, namely when any two of the variables LRT, forgetting, R-repetition and unique response were paired, with the single exception of the LRT–R-repetition relationship. Note that the focal position occupied by unique response for N-words was shared with forgetting in the case of E-words.

Significant correlation without significant association. In these cases, the correlation cannot be explained in the direct manner of the previous pattern. Instead, it must be accounted for as a side-effect of other associations relating each of the correlated indicators to a third. Thus, for both N-words and E-words, LRT and R-repetition were correlated, although R-repetitions themselves did not have LRTs. For N-words, this correlation was mediated by the unique response variable. Unique responses had an above-chance probability of being R-repetitions; they also tended to require LRTs. Consequently, stimuli eliciting many unique responses tended to amass many R-repetitions and also many LRTs, albeit independently: hence the significant LRT–R-repetition correlation. In the case of E-words, both unique response and forgetting were in a position to mediate the correlation between LRT and R-repetition in this way. This particular instance of correlation without association was the only one among E-words, but there were seven such cases for N-words. All the significant correlations between pairs of emotional indicators for N-words appear to have depended on the mediation of the (non-emotional) unique response variable, which created secondary connexions between indicators which were not primarily associated.

Significant association without significant correlation. This is a more difficult outcome to interpret. There were significant associations between unique response and S-repetition for both N-words and E-words. In the former case, this gave rise to a significant correlation, in the latter case, not: why? Again, among E-words, R-repetitions suffered from both forgetting and misremembering, but the forgetting–R-repetition correlation was significant while the misremembering–R-repetition correlation was not: again, why? Two general ways in which such discrepancies might arise will be described. First, suppose one of the indicators to have an approximately rectangular distribution over stimulus words. No measure so distributed could be expected to correlate with another variable, although ample opportunity for truly intrastimulus associations would remain. But the cases observed here, such as the absence of correlation between unique response and S-repetition for E-words, do not

fit this pattern. A second explanation may begin with the observation that, in this particular case, there were many more unique responses than S-repetitions in the data. Even if all S-repetitions had been unique responses (and they were not), there would still have remained a large number of unique responses which might have arisen in other ways. In fact, a good many of these surplus unique responses appear to have involved LRTs, for there was a significant association between LRT and unique response, although LRT and S-repetition were not associated. There is a hint here of a possible subdivision of unique responses into two distinguishable types, one arising in such a way as to facilitate S-repetitions, the other so as to throw up LRTs. (A clear-cut dichotomy would require a significant dissociation of S-repetition from LRT, but this was not found.) The score for unique response given to each stimulus-word sums both varieties of unique response together indiscriminately. Unless the two varieties were very similarly distributed over stimulus-words, the unique response scores used in the correlations would not reflect the distribution of either type very well. In this case, the unique response measure appears to have been closer to the variety associated with LRT than to that associated with S-repetition, since the correlation between LRT and unique response was significant. This, or something like it, may account for the non-significance of the correlation between S-repetition and unique response. In more general terms, different associative processes may underlie different instances of the same indicator, and these alternative processes may be differentially conducive to the other indicators; in such a situation, the interstimulus scores on the former will be multiply determined, and may not correlate with the corresponding scores on the latter.

Discussion

The attempt to observe the effects of associative difficulty by analysing responses to a set of N-words may be counted successful. It has been shown that all five so-called emotional indicators are sensitive to associative difficulty, and that this can and does lead to significant interstimulus correlations between pairs of indicators not directly associated with one another. With the single exception of the significant association between S-repetition and misremembering, the entire body of findings relating to N-words can be simply and parsimoniously described in terms of the operation of this one determinant of associative difficulty.

The analysis of responses to a set of E-words was intended to illuminate the effects of emotional disturbance as a second determinant of word-association performance. It seems clear that the emotional factor did have an effect, as comparisons of the two correlation matrices and the two co-occurrence patterns suggest. It is less easy to say precisely what the effects of emotion were. From the correlations alone, it might be deduced that only S-repetition and misremembering, which between them were party to all the non-significant correlations, are sensitive to emotion, while the remainder continue to reflect associative difficulty. But this interpretation will not do. Some of the correlations between emotional indicators (e.g. between LRT and forgetting) which were significant in both matrices, arose in demonstrably different ways in the two cases (reflecting association in E-words, but only the mediation of unique responses in N-words). The emotional determinant produced associations between

certain pairs of indicators which were not present in the responses to N-words. At the same time, associative difficulty remained an important factor in associating to E-words, and all the emotional indicators studied remained strongly associated with it.

Which single indicator provides the best evidence of emotional disturbance? This question highlights dramatically the need for a response-specific criterion of emotion. In the absence of such a criterion, it is impossible to say which indicator is most closely associated with emotional disturbance. The emotionality ratings of the stimulus words, however, provide a criterion at the interstimulus level, and so the mean ratings of the sixty stimuli were correlated with the frequencies of the six indicators (see Table 3). The zero correlation of emotionality with unique response was built into the experiment by the selection of stimuli, but the almost equally low correlation with LRT does nothing to enhance that variable's status as an emotional indicator. Only forgetting and R-repetition emerged as significant correlates of the emotionality ratings, and so it may be suggested that these two emotional indicators are more sensitive to emotion than the other three studied.

Table 3. *Correlations between frequencies of indicators to each stimulus and mean stimulus emotionality ratings*

Indicator	Unique R	LRT	R-repetition	S-repetition[1]	Forgetting	Misremembering
r	-0.05	0.09	0.35*	0.21	0.40*	0.25

[1] Biserial correlation; others are product-moment correlations, with 58 D.F.
* $P < 0.01$.

The finding that forgetting was the single most valid emotional indicator is in accord with previous views (Smith, 1922; Rapaport *et al.* 1946; Levinger & Clark, 1961). Hull & Lugoff (1921) also gave it some weight, but they put forward repeating the stimulus before responding and misunderstanding the stimulus as the two best indicators. Neither of these occurred with sufficient frequency to warrant statistical analysis in the present investigation. This may have been due to the use of visual rather than auditory presentation of the stimuli; it may be easier to mishear than to misread, and 'echoic' responses are probably more strongly elicited by sound than by sight. The two types of reproduction failure, forgetting and misremembering, do not appear to be functionally equivalent. Only the former seems to carry much emotional significance. Both were sensitive to associative difficulty, however, and it is interesting to note that the two correlated highly for N-words, when both were indicating associative difficulty alone, but not for E-words, where forgetting was determined by both associative difficulty and emotional disturbance.

The finding that R-repetition is a good emotional indicator conflicts with the view of Hull & Lugoff (1921), who, finding it had fewer co-occurrences with LRT than would be expected by chance, concluded that it was not criterial of emotion. Even if LRT were a good guide as to what is and what is not indicative of emotion, there would appear to be a misinterpretation here. As Rapaport *et al.* (1946) have pointed out, emotional indicators 'may *replace* as well as *accompany* each other' (p. 35), and such could be the case with LRT and R-repetition. R-repetition affords 'a quick

way out' of the problem situation created by the emotional stimulus, while other solutions require more time; both lead to responses which are subject to forgetting on the retest. In more general terms, a significant dissociation between two emotional indicators at the response level could well accompany a significant interstimulus correlation between them. In such a case, the two indicators would be alternative symptoms with the same aetiological implications.

It may seem odd that, in the E-words correlation matrix, the two better emotional indicators, forgetting and R-repetition, should have correlated so highly with the non-emotional unique response variable and with the poorest of the emotional indicators, LRT. It must be remembered that all the indicators continued to reflect associative difficulty even when emotion was present. The observed patterns of correlation and association are necessarily influenced by the nature of the sample of data on which they are based. In the present case, both samples of data encompassed a wide range of associative difficulty levels. In one sample (the data relating to N-words) the emotional factor was held constant at a minimal level; under these conditions, all the emotional indicators showed striking relationships with associative difficulty. If data were available for a set of stimulus-words in which associative difficulty was held constant at a low level, while all levels of stimulus emotionality were represented, it would not be surprising if all the indicators were found to be highly correlated with emotional disturbance. The present E-word stimulus list was not of this type, since the lower emotionality levels were unrepresented, and a full range of associative difficulty levels was retained. It is not surprising, then, that the correlation matrix demonstrated the capacity of the indicators to diagnose associative difficulty fully as much as their capacity to indicate emotional disturbance.

Would pairs of indicators provide better criteria for emotional disturbance than single indicators? Hull & Lugoff (1921) concluded that 'two indicators are distinctly more significant than one' (p. 136). The evidence of the present study provides partial confirmation for this view; four of the nine pairs of emotional indicators showed significant tendencies to co-occur among responses to E-words. The other five indicator pairs, presumably, are unlikely to discriminate emotional disturbances from non-emotional difficulties. It may be doubted, however, whether the same pairs of indicators would always provide the best discrimination. Further evidence, from studies with other word lists, should be collected before any attempt is made to generalize about the validity of specific indicator pairs.

A final point of some importance concerns the problem of using the word-association procedure to select emotional stimulus material for experiments on perceptual defence, repression, etc. The common practice of selecting words with LRTs has little to recommend it except its ready availability. Other single indicators, such as forgetting, would be more satisfactory criteria, and specified pairs of indicators might prove even more acceptable. But there is a severely practical objection to replacing LRT by one of these alternatives. Every response has its reaction-time, and even when a subject displays very quick reactions the experimenter can still pick out those responses which had relatively long RTs for that subject. But if a forgetting criterion is employed, a subject who had perfect recall on the retest can only be dismissed from the experiment. Apart from LRT and GSR, which seems equally poor as an emotional indicator (Smith, 1922; Levinger & Clark, 1961), all

the emotional indicators are dichotomous variables, and most of them are of fairly infrequent occurrence. Experiments involving these indicators as criteria for selecting enough emotional stimuli for each subject to make experimentation worth while would certainly create a high wastage rate among subjects; using pairs of these indicators as selection criteria would prove even more impracticable. When completed, such experiments would be studies of the reaction of unrepresentative groups. Perhaps we ought to experiment with departures from current word-association practices, either by establishing special stimulus lists or by manipulating the instructions in some way, which would increase the frequency of the better emotional indicators without impairing their validity. But this may be asking for the moon.

The data were collected when the author was a Carnegie Research Student in the University of Aberdeen. They were analysed at Bedford College, University of London. A preliminary report of this research was read as a paper to the British Psychological Society at its London Conference in December 1963.

References

Attneave, F. (1959). *Applications of Information Theory to Psychology*. New York: Holt.

Brown, W. P. (1961). Conceptions of perceptual defence. *Brit. J. Psychol. Monogr. Suppl.* no. 35.

Copeland, H. A. (1932). Familiarity of the stimulus in the word association test for measuring emotions. *J. abnorm. soc. Psychol.* **27**, 89–91.

Hubbard, L. M. (1924). Complex signs in diagnostic free association. *J. exp. Psychol.* **7**, 342–57.

Hull, C. L. & Lugoff, L. S. (1921). Complex signs in diagnostic free association. *J. exp. Psychol.* **4**, 111–36.

Laffal, J. (1955). Response faults in word association as a function of response entropy. *J. abnorm. soc. Psychol.* **50**, 265–70.

Levinger, G. & Clark, J. (1961). Emotional factors in the forgetting of word associations. *J. abnorm. soc. Psychol.* **62**, 99–105.

McClatchy, V. R. & Cooper, M. (1924). A psychological study of linguistic abilities with reference to the results of word association tests. *J. exp. Psychol.* **7**, 371–81.

Rapaport, D. (1942). *Emotions and Memory*. Baltimore: Williams & Wilkins.

Rapaport, D., Gill, M. & Schafer, R. (1946). *Diagnostic Psychological Testing*, vol. II Chicago: Year Book Publishers.

Schlosberg, H. & Heineman, C. (1950). The relationship between two measures of response strength. *J. exp. Psychol.* **40**, 235–47.

Smith, W. W. (1922). *The Measurement of Emotion*. London: Kegan Paul.

Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Bureau of Publications, Teachers College, Columbia University.

(*Manuscript received 27 April 1965*)

Brit. J. Psychol. (1965), 56, 4, *pp.* 413–419
Printed in Great Britain

EFFECTS OF A SUBSIDIARY TASK ON SHORT-TERM MEMORY

By BENNET B. MURDOCK, JR.
*Department of Psychology, University of Missouri**

If there is a limited-capacity mechanism in STM then introducing a concurrent subsidiary task should adversely affect recall. Two experiments on free recall were conducted with card sorting as the subsidiary task. In the first experiment subjects dealt cards into one pile, into two piles by colour, or into four piles by suit while lists of common English words were being read. Subjects sorted cards only during presentation of the lists. As the subsidiary task became more demanding the number of words correctly recalled decreased. In the second experiment sorting by suit was combined with free recall, and the payoffs (relative importance of the two tasks) were varied. Performance on both the recall and the card sorting tasks deteriorated as the other task was stressed. Differences in recall could not easily be attributed to differences in original learning, and the results suggested that the subsidiary task interfered with rehearsal and/or decreased total presentation time for free recall.

The results of two recent experiments on short-term memory (STM) seemed to suggest a limited-capacity mechanism of the type suggested by Broadbent (1958). In the first experiment (Murdock, 1964) the serial position curve for paired-associate retention changed with practice, but the decrease in primacy was balanced by an increase in recency. In the second experiment (Murdock, 1965) it was found that, holding total presentation time constant, one could trade off repetition and presentation time per pair.

In both experiments it seemed as if the limitation might be on the rate at which information could be processed since, with total time constant, a fixed amount could be recalled. If there really is a limited-capacity mechanism operating in STM, then it should be possible to decrease retention by introducing a subsidiary task. The use of a subsidiary task is best illustrated by a recent experiment of Broadbent & Heron (1962). There, subjects were given a digit cancellation task concurrently with a modified memory-span task. In effect, subjects had to listen to spoken letters of the alphabet and report the one letter in ten that was repeated while at the same time crossing out digits with the target digit varying haphazardly. The former was considered to be the subsidiary task; performance of the primary task (digit cancellation) varied both with the memory load required (manipulated experimentally) and the age of subjects. As the authors say, it would seem that, '. . . immediate memory involves continuous action on the part of some mechanism which is also needed for perception, and that this mechanism cannot handle both tasks simultaneously' (p. 196).

In the present study two experiments were conducted as a further test of the limited capacity hypothesis. In both cases free recall was used as the STM task while card sorting was used during list presentation as the subsidiary task. In the first experiment there were three different types of card sorting, and it was hypothesized that as the card sorting became more demanding recall should deteriorate. In the

* Now at Department of Psychology, University of Toronto.

second experiment the 'payoffs' (or 'instructions', see Edwards, 1961) were varied to see if the relative performances on the primary and subsidiary tasks could be changed.

EXPERIMENT I

Method

Experimental tasks. For free recall all lists consisted of twenty common English words. Specifically, they were two-syllable words taken from the Thorndike & Lorge (1944) list with *G*-count of 20 or above; homonyms, contractions, and archaic words were excluded. All lists were recorded on magnetic tape and played back on an 'Ampex F-4460 Fine Line' recorder. Presentation rate was 60 words/min, obtained by synchronizing with an electric metronome beating once a second. Each list was presented only once; at its conclusion the subject wrote down as many words as he could remember without regard to order.

There were three types of card sorting: plain sorting, sorting by colour, and sorting by suit. Ordinary playing cards were used. For plain sorting, the subject merely dealt out as many cards as possible one at a time. For sorting by colour, red cards were placed in a pile to the left, black cards to the right. For sorting by suit, the arrangement was diamonds, spades, clubs, and hearts reading from left to right, respectively. A relatively incompatible ordering of suits was used in an attempt to maximize the difficulty of the task.

The two tasks were either performed separately or together. When together, the signal to start card sorting was given 2 sec before the first word in the list while the signal to stop sorting was given 2 sec after the last word in the list. When sorting alone the same duration (i.e. 23 sec) was employed. The amount of time allowed for the recall of each list was 90 sec, and there was a short ready period before each list and/or card-sorting trial.

Subjects. There were thirty-four subjects, undergraduate students of both sexes who were fulfilling a course requirement.

Procedure. The first five trials were practice trials with the tasks given separately. Of these five the first three were plain sorting, sorting by colour and sorting by suit (in that order). The last two were free recall trials without card sorting. Then came nine test trials with card sorting and free recall combined. Finally, to provide a before–after comparison, five additional trials followed with either card sorting or free recall; the order here was the same as in the initial practice trials.

For the nine test trials each block of three contained one trial with each of the three sorting conditions (i.e. plain, colour, or suit). Further, counterbalancing with a Latin Square was used to ensure that (*a*) each list occurred equally often with each type of card sorting, and (*b*) each type of card sorting occurred equally often as the first, second, and third trial within each block.

All subjects were tested individually and the entire experiment required about 40 min.

Results

The mean number of correct recalls on the first two (practice) lists were 7·3 and 7·5 words, respectively. On the basis of previous results (Murdock, 1960) the predicted value for a 20-word list presented at a rate of 60 words/min would be 7·3 words. The agreement seems close enough to suggest that the sample (of both subjects and lists) was reasonably representative.

For the test trials the mean number of correct recalls was 5·82, 4·65, and 4·35 words for plain, colour, and suit sorting, respectively. An analysis of variance gave an *F*-ratio of 37·2, which was statistically significant at well beyond the 0·001 level. Both adjacent comparisons (i.e. plain *v.* colour and colour *v.* suit) were significant by sign tests at the 0·01 level (one tail), so it would seem as if the subsidiary task had the expected effect. Thus, the more taxing the subsidiary task the poorer the recall.

On the subsidiary task as would be expected, all subjects sorted most cards with plain sorting and the fewest cards with suit sorting. Also of course there was a practice

effect; when sorting alone all subjects did better after than before (i.e., making the before–after comparison of the experimental design). If one takes as a base-line the 'after' condition (so that the confounding with practice effects works to minimize the difference) the free recall task reduced the number of cards sorted by about 20 %. Since again all thirty-four subjects showed this effect there can be little doubt that the inhibitory effect is mutual; card sorting reduced recall and the free-recall task reduced the number of cards sorted.

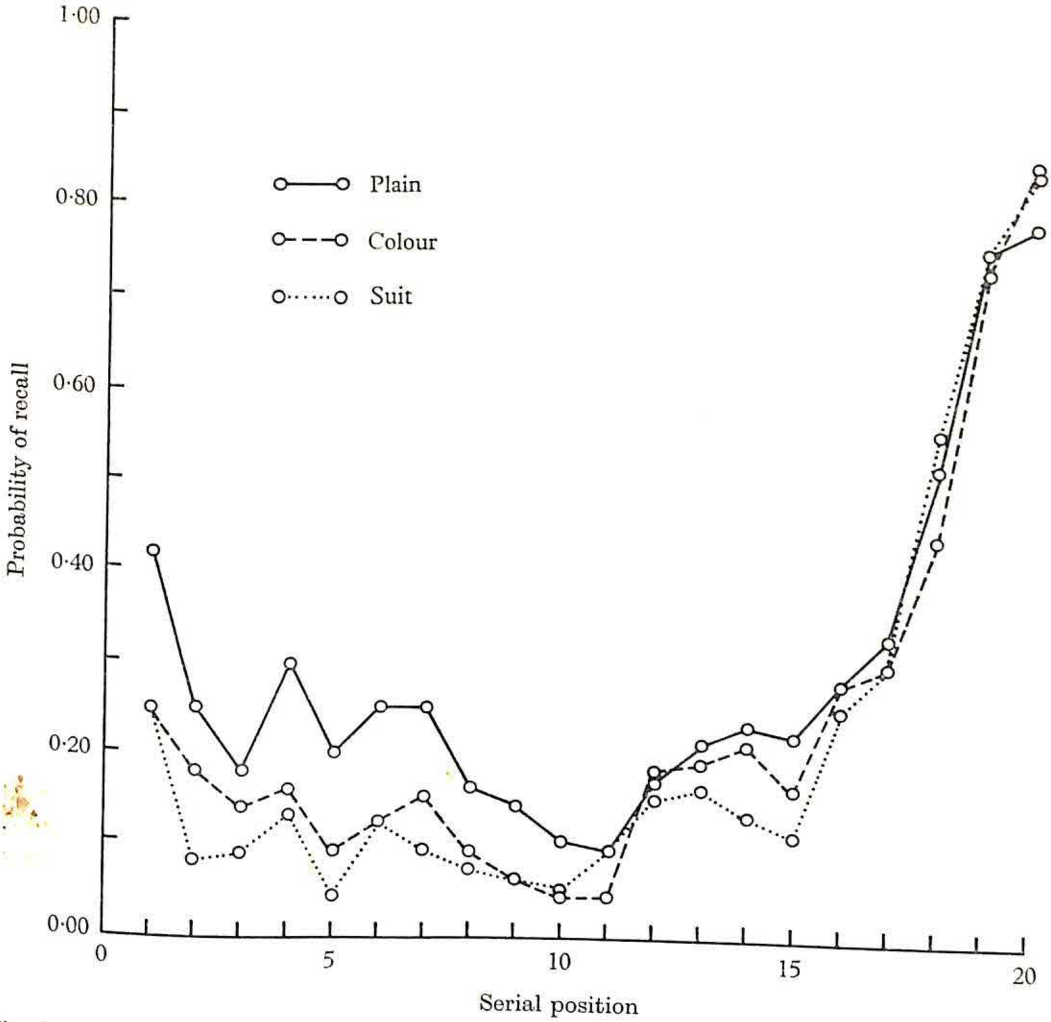


Fig. 1. Serial position curves for the recall of items out of lists of twenty common words which had been presented at 1/sec while subjects carried out one of three subsidiary card-sorting tasks: plain (dealing into a single pile), sorting by colour or sorting by suit.

As is generally true in free-recall experiments, extra-list intrusions were relatively infrequent. A count for the test lists (i.e. free recall with card sorting) showed that the mean number of intrusions per list was 0·75, 0·76 and 0·79 words for plain, colour and suit sorting, respectively; an analysis of variance gave an F-ratio less than unity. On the two practice lists (without card sorting) the mean number of intrusions per list was 0·87 words. Thus, there is no indication that the subsidiary task had any effect whatsoever on extra-list intrusions.

Finally, the serial position curves for the three conditions are shown in Fig. 1. Each point is based on the combined performance of all thirty-four subjects on the three test lists of each condition. In general each of the three serial position curves shows the typical effect: marked recency, slight primacy, and a level middle section (see Murdock, 1962). However, further consideration of the significance of these curves will be postponed to the concluding discussion.

EXPERIMENT II

As a further test of the limited-capacity hypothesis, this experiment was conducted to determine if performance could be affected by varying the 'payoffs'. Card sorting and free recall were combined on all trials; all that was done was to try to vary the payoffs by the rather crude (but simple and easily understood) expedient of instructing subjects to 'concentrate primarily on' one task or the other. Justification for manipulating payoffs by means of instructions can be found in Edwards (1961).

Method

Experimental tasks. The free-recall task was identical with that of Expt. I in all respects except that a different sample of words was drawn from the Thorndike–Lorge set. The words had the same characteristics, however; G-count of 20 and up, two syllables in length, etc. For card sorting only sorting by suit was used; the same (relatively incompatible) ordering of diamonds, spades, clubs, and hearts was used.

Procedure and subjects. All subjects were given twelve trials, and both card sorting and free recall were used on all twelve trials. The instructions were typed out for the subject to read and (in part) read as follows: 'During each list you will have to sort cards as well as listen to the list. However, what we are mainly interested in finding out is how well people can emphasize one task or the other. Therefore, before each list starts the experimenter will tell you whether you should emphasize CARD SORTING or FREE RECALL. If the former, do your best to sort as many cards as possible and don't worry how many words you can remember. In other words, concentrate on the card sorting. If the emphasis is on FREE RECALL, do your best to remember as many words as possible and don't worry about how many cards you sort. In other words, concentrate on learning the list.'

After the subject had read these instructions he had to summarize them in his own words so that the experimenter could be sure they were understood.

The two experimental conditions (i.e. emphasis on card sorting or free recall) alternated in an ABBA manner and half the subjects started with each condition. Thus, each list occurred equally often under the two conditions. On each list the subject started card sorting as soon as he heard the first word in the list and was told by the experimenter to stop sorting on the first beat of the metronome after the last word in the list.

There were twenty-three subjects all from the same population as before. All subjects were tested individually, and there were no practice lists. In all other respects (for example, length of the recall period) the procedure was the same as that in Expt. I.

Results

The mean number of correct recalls under the two conditions were 5·12 and 4·09 words for emphasis on free recall and card sorting, respectively. Since there were no subjects who recalled more under the latter condition and only three ties (equal recall under the two conditions), the 20–0 split is quite significant by a sign test.

For card sorting, the mean numbers of cards correctly sorted were 16·6 and 23·0, respectively. (In both experiments errors in card sorting were so infrequent that no analysis was attempted.) Since all twenty-three subjects showed this effect, again the statistical significance is beyond question. To compare relative decrements, free recall under card-sorting emphasis dropped to 80 % of its normal (emphasized) value while card sorting with emphasis on free recall dropped to 72 % of its normal value. To a first approximation the decrements appear to be equal.

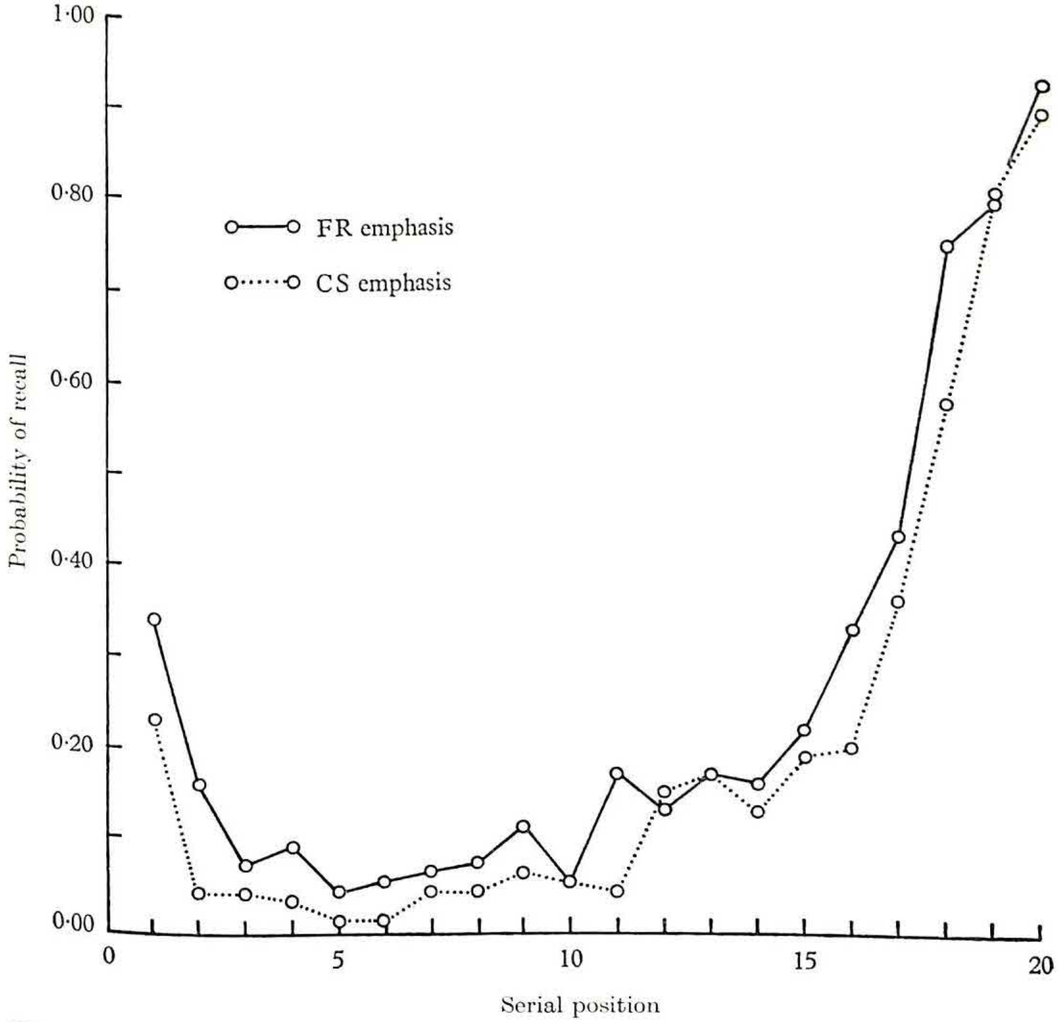


Fig. 2. Serial position curves for the recall of items out of lists of twenty common words which had been presented at 1/sec while subjects sorted playing cards into suits under instructions which emphasized either speed of sorting the cards (CS) or the free recall of as many words as possible (FR).

Extra-list intrusions were again rather infrequent; and again there was no difference between conditions. The mean numbers of intrusions per list were 1·26 and 1·14 for emphasis on free recall and on card sorting, respectively; the 10–9 split (with 4 ties) was of course not significant. The serial position curves for the two conditions are shown in Fig. 2; again there is the marked recency and slight primacy effect.

DISCUSSION

The results were clearly confirmatory; card sorting depressed recall, the more taxing the card sorting the poorer the recall, and by varying the payoffs one can affect the relative performance on the two tasks. Thus, as was shown by Broadbent & Heron (1962), use of a subsidiary task has a clear-cut effect on STM.

In any experiment of this type it is clearly desirable to try to separate out the effects of acquisition and retention (Underwood, 1964). Inspection of the two serial position curves clearly suggests that group differences cannot easily be attributed to differences in degree of original learning. That is, under certain conditions (for example, when there is no ceiling effect) recall scores for the last few items in a list may be taken as a measure of degree of original learning in studies of STM. In Expt. I the only visible difference in performance over the last four serial positions was in the wrong direction (i.e. at serial position 18 recall with colour sorting was lower than the other two conditions); otherwise the curves appear to be indistinguishable. In Expt. II the two curves were again indistinguishable for the last two serial positions; only at the earlier position did they diverge. Thus, it may be suggested that degree of original learning was not adversely affected by the subsidiary task.

To what are the effects due? One possibility is of course that the subsidiary task prevents rehearsal (or, in Broadbent's model, recirculation of items in the short-term store) and the rehearsal (particularly of early items) is necessary to prevent decay. The more demanding the subsidiary task or the more attention given it the more effectively is rehearsal prevented. The fact that there are fairly marked differences in the front portion of the serial position curves (i.e. primacy effect) is not inconsistent with this notion. On the other hand, it is not necessary to conclude that in fact the effect is on storage; perhaps retrieval is more difficult because variations in the nature of the subsidiary task affect the amount of interpolated interference.

Another possibility, not incompatible with the first, is that the subsidiary task simply decreases the available presentation time for free recall. Previously it has been shown (Murdock, 1960) that free recall after a single presentation is a linear function of total presentation time, and it seems not unreasonable that the subsidiary task decreases the total learning time. As one bit of indirect evidence, in Expt. I the rate of gain for card sorting (colour and suit only) was essentially the same (1·8 and 1·5 bits/sec) with and without the free-recall task. If the information-processing component of card sorting is independent of free recall then perhaps the converse is true as well. If so, perhaps some sort of time-division multiplexing system is involved.

In any case, the results of the present experiments seem to show quite clearly that a subsidiary task can have an adverse effect on STM. While the formal similarity between the subsidiary task and the free-recall task would seem to be low, the possibility that the subsidiary task required (potentially interfering) verbal mediating processes cannot of course be excluded. Finally, the use of a subsidiary task may be a useful technique in studies of STM. Instead of a subsequent task designed to prevent rehearsal of prior material (e.g. Brown, 1958; Peterson & Peterson, 1959) perhaps two concurrent tasks with low formal similarity are a better solution to the rehearsal problem.

This investigation was supported by PHS Research Grant MH-10,882 from the National Institute of Mental Health, Public Health Service, U.S.A. The author gratefully acknowledges the assistance of Mrs Anne Stricklin in testing subjects and analysing data.

References

BROADBENT, D. E. (1958). *Perception and Communication*. New York: Pergamon Press.

BROADBENT, D. E. & HERON, A. (1962). Effects of a subsidiary task on performance involving immediate memory by younger and older men. *Brit. J. Psychol.* **53**, 189–98.

BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. exp. Psychol.* **10**, 12–21.

EDWARDS, W. (1961). Costs and payoffs are instructions. *Psychol. Rev.* **68**, 275–84.

MURDOCK, B. B., JR. (1960). The immediate retention of unrelated words. *J. exp. Psychol.* **60**, 222–34.

MURDOCK, B. B., JR. (1962). The serial position effect of free recall. *J. exp. Psychol.* **64**, 482–8.

MURDOCK, B. B., JR. (1964). Proactive inhibition in short-term memory. *J. exp. Psychol.* **68**, 184–9.

MURDOCK, B. B., JR. (1965). A test of the 'limited capacity' hypothesis. *J. exp. Psychol.* **69**, 237–40.

PETERSON, L. R. & PETERSON, M. J. (1959). Short-term retention of individual verbal items. *J. exp. Psychol.* **58**, 193–8.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Teachers College, Columbia University.

UNDERWOOD, B. J. (1964). Degree of learning and the measurement of forgetting. *J. verb. Learn. verb. Behav.* **3**, 112–29.

(*Manuscript received* 16 *January* 1965)

Brit. J. Psychol. (1965), **56**, 4, *pp.* 421–430
Printed in Great Britain

TWO ARTISTS WITH PROTAN COLOUR VISION DEFECTS

BY R. W. PICKFORD
Psychology Department, Glasgow University

Two artists, one a protanope and the other protanomalous, were given tests of colour vision, including anomaloscope tests. Both knew that they had some kind of colour vision defect, but both considered that it had not affected their art to any extent. The protanomalous artist liked to make golden yellow, blue and red contrasts in his work, but he seldom used greens. The protanope usually made fawn, yellow, orange and blue harmonies, from which reds as seen by the normal eye were excluded, and he used greens where appropriate. Both artists may be said to have exploited contrasts and harmonies of colour which made their work acceptable to the normal as well as to themselves.

In a previous paper the writer (Pickford, 1964) reported on the colour vision and the paintings of a deuteranomalous artist. The present paper deals with two other artists, an amateur and a professional, who had protan colour vision defects, the amateur artist, Mr Pratul A. Ray, being protanomalous, and the professional being a protanope. The writer would like to record his appreciation of the trouble they took in carrying out tests and answering questions, and their courage in allowing him to publish this paper. If an artist could be found who was a deuteranope, it would complete the series, except for rare defects.

The writer has given an account of the colour mistakes made by a deuteranope in copying in colour a painting of a woman and a cat sitting by a fire (Pickford, 1951, p. 233). His mistakes, such as making the yellowish edges of the flames lime green, a dull red flower-pot green, and the violet sleeves of an apron bright blue, are the kinds of mistakes which a deuteranope would be expected to make. He was not, however, an artist, except in the sense of taking an interest in copying paintings. The present paper will show that if a colour-blind or colour-defective person is to become a successful artist, either in an amateur or a professional way, he must acquire, consciously or unconsciously, a technique for adapting his choice of colours so that it fits in both with his own and with normal people's colour vision, and does not involve downright confusions of colour made unintentionally.

In the previous paper (Pickford, 1964) the writer referred to works dealing with the effects of defective colour vision in art, and these should be mentioned again here (Collins, 1925, frontispiece and pp. 171–82; Pickford, 1951, pp. 197–8; Riddell, 1949; Trevor-Roper, 1959). In addition, Kalmus (1965) has given a useful account of the problems, and Strebel (1933) and Pickford (1965) should be consulted.

A PROTANOMALOUS ARTIST

Mr Pratul A. Ray, an Indian teacher in a special school for delinquent boys in Scotland, aged 36, said that he was colour blind, and agreed to be tested. He is an amateur artist of some ability, has exhibited paintings in India and Scotland, and is particularly successful at running a voluntary art club for the boys in the special school

in which he works. He came to the Psychology Department, Glasgow University, where he was tested by Dr A. Adam, of the Tel-Hashomer Government Hospital, Tel-a-Viv, Israel, who was then on a visit sponsored by the World Health Organization, to research on colour blindness and its inheritance.

The subject's colour vision

The following tests were carried out by Dr Adam, with the results shown, but the Nagel anomaloscope test was carried out by the writer.

(1) *Ishihara Test.* In this test he made the following readings, the correct reading being given in brackets for comparison:

Plates *1–5*: (12) 12; (8) 3; (6) 5, 8; (29) 70; (57) 67;
Plates *6–9*: (5) 2; (3) 5; (15) 17; (74) 21;
Plates *10–13*: (2) —; (6) —; (97) —; (45) —;
Plates *14–17*: (5) —; (7) —; (16) —; (73) —;
Plates *18–21*: (none) 5?; (none) —; (none) —; (none) —;
Plates *22–25*: (26) 26; (42) —2; (35) —5; (96) —.

This amounted to twenty plates wrongly read in twenty-four, the first of the twenty-five being a 'joker' readable by everybody. Owing to the facts that he read Plate *22* correctly, but saw the right hand, or purple figures, in Plates *23*, *24* and *25* rather than the left hand, or red figures, he would be classed as a protanope, or protanomalous, on this test.

(2) *Farnsworth Dichotomous Test* (Panel D-15). This test he did correctly twice, but slowly.

(3) *Nagel anomaloscope.* On this particular instrument in the Psychological Laboratory he obtained the results shown in Table 1. The test was carried out by the tester changing the setting according to the method of limits, and shows the subject clearly to be protanomalous, of the simple type (*PA*) but with a rather large matching range. There was a deviation of about 23 units from the normal to the red side of the equation, with a darkening of the red from about 20 units, at a scale reading of 60, down to about 5 units. In the green–blue test he had an enlarged matching range. The fact that such a subject could pass the Farnsworth Dichotomous Test, even with some difficulty, shows that it is not a satisfactory test.

Table 1. *Results of the Nagel anomaloscope test for the Indian artist, showing him to be protanomalous*

	Matching range	Mid-point	Brightness
Red–green	62–64	63	5
Normal	39·5–40·5	40	20
Green–blue	5–11	8	19 ± 7
Normal	6·5–7·5	7	20

(4) *Pickford–Nicolson anomaloscope.* With this instrument (Pickford & Lakowski, 1961) there are three tests. In the first test a mixture of red and green lights is matched against a yellow standard; in the second a mixture of green and blue is matched against a blue-green standard; and in the third a mixture of blue and yellow is matched against a very desaturated pinkish standard.

On this instrument he obtained the results shown in Table 2.

This test confirmed the diagnosis of *PA* with a rather large matching range, as the usual matching point on this particular instrument for normal subjects is 25 units, with a matching range of about one unit, and the brightness is about 20 units instead of 11 units at a *RG* scale reading of 11 units. His results were somewhat variable.

On this anomaloscope he also did the blue–green test and the yellow–blue tests, with the results shown in Table 3. In these tests he also showed small deviations of mid-

Table 2. *The Indian artist's results on the Pickford–Nicolson anomaloscope: red–green test*

Trial	Matching range Red–green	Mid-point	Brightness range
1	10–27	18·5	10–27
2	11–16	13·5	11–16
3	11–22	16·5	11–20
4	11–25	18·0	11–24
Normal	±0·5	25	11±0·5

point (towards green and towards yellow), considerable increases of matching range, and a loss of brightness towards yellow. Thus he is not only protanomalous (*PA*) but also green–blue and yellow–blue weak.

Table 3. *The Indian artist's results on the Pickford–Nicolson anomaloscope: green–blue and yellow–blue tests*

	Matching range	Mid-point	Brightness mid-point
Green–blue	39–49	44	62
Normal	±0·5	44	62
Yellow–blue	16–27	21·5	32
Normal	±0·5	20	45

This subject's colour vision differs from the normal in the following ways: (*a*) the red end of the spectrum is greatly darkened for him; (*b*) he has some confusion of yellow with green, and, when sufficiently darkened, with red; (*c*) desaturated blue and green are to some extent confused; and both yellow, when sufficiently darkened, and blue tend to be confused to a certain extent with greys. At some times, he is almost weak enough on red and green to be called extreme protanomalous (*EPA*) but he is definitely not dichromatic.

The subject's paintings

Mr Ray's art is very characteristic. He uses saturated reds, blues and golden yellows to make vivid contrasts, and his paintings are powerful and dynamic. They are not representational, and he does not paint to imitate perceived objects or scenes, but, as explained in the notes given below, to express his inner feelings. He rarely uses green. In a conversation with the writer recorded in shorthand by Miss B. J. Reid, Secretary of the Psychology Department, Mr Ray has explained some of his attitudes to his colour vision defect and to his art. The conversation is given below.

Answers by Mr Ray to questions about his art

Did you know you were colour blind when you did the paintings?	Yes.
When did you find out?	1947.
How?	I went for my Army Commission and was given a colour vision test and failed that test.
Was this a great surprise?	Yes. A big surprise, as I did not know anything about having defective colour vision at all then.
Did it make you think about art in a different way?	No.
Can you tell us whether your colour vision has any influence on your choice of colours in painting?	Possibly, but I am not aware of it when painting.
You know that you use red, gold and blue in the main, and seldom green?	Yes.
Did you think this was due to your colour blindness?	It could be, but as I said, I am not aware of the fact that I am colour blind while I paint. It never crosses my mind.
Does red seem to have a strong quality for you?	Yes, to me! Now that I know I am colour blind, I feel it must be quite different from the normal standard.
It could be adequate when you use very bold colours as you always do?	Yes.
Is this due to over-compensation for a deficiency?	Possibly. Again, not planned.
As for your present attitude with your present knowledge that you are a defective in colour vision, do you think your feelings about your paintings will be changed?	I do not think so. No. I might, perhaps, rationalize, but that I guess would be done quite unconsciously. In any case, I am not easily influenced by the so-called standards of normality as far as art is concerned.
Is there anything of a general kind that you could tell me about your use of colours?	When I am painting, and if I am in a gay mood, I am inclined to use bright contrasting colours and the lines I draw have a definiteness to reflect the natural exuberance I feel then. I like contrasting them and at the same time planning them for an overall soothing effect. The opposite of this statement is also applicable, depending on the mood.
In your contrasts, you never choose green to contrast with red.	Very seldom. It is usually yellow I choose, or blue to contrast with my red.
Is this because you did not like green?	I am very fond of the colour green.
None of your paintings are realistic, and your colours in them always reflect your mood rather than the real appearance?	Yes. Quite true. Most of my paintings have the characteristics of an introvert. They are essentially inward looking. They are part of a 'know thyself' process.
Tell us more about how the knowledge of your colour blindness affected you.	The knowledge that I was colour blind never came as a shock. It was a surprise without shocking me. For a long time after this I preferred to take black and white photographs, my first love in the field of creative activity, rather than coloured ones. I was not afraid of colour, but I found I could express myself better in black and white.

Did you never make black and white paintings?	No. But I have started taking colour photographs seriously and I quite enjoy taking coloured transparencies now.
Do you think your colour blindness has any effect on your appreciation of other people's art?	No. As a matter of fact I think it helps me to appreciate the modern painters much more than the others. Some of them must be colour blind by design! I do not get a shock when I see violent outbursts of colour on a canvas. They sometimes provide a sense of camaraderie!
Perhaps you do not see the outbursts of colour!	I see them, but not necessarily what a normal person would see on a canvas.
Does your wife like your paintings?	My wife always says she likes them, but does not want to live with them. I have been trying to find out why. The reason must be that they are too exuberant and too violent.
Can you give any examples of difficulty with colours in daily life?	Yes! At a cricket match the score board is black and they put on a red light to indicate the bowler. I could not find this red light, and always had to ask who was the bowler. I had no difficulty with traffic lights.

A protanopic artist

A professional artist, who said in general conversation that red had a different scale of values for him from what it had for most people, was invited to undertake tests of colour vision.

He wrote: 'as a rule large areas of colour give me no trouble but small dots or narrow stripes confuse me'.

In his paintings, both oil and water colour, he uses many colours like ochres, browns and blues, but trees and other objects which he knows to be green he paints in green. When 'red' (for him) is introduced it is normally a rather dark orange, and he hardly uses real reds at all. Thus his paintings may be generally quite true to nature if he chooses subjects like a Majorca sea-and-landscape, in which blues appear for sea and sky; brown, grey and fawn for the hillsides; and often there is no red, and little green except for a few olive trees on the hills. Purple and violet are usually excluded. Although green is used when appropriate, it is sometimes crude and harsh. There is no tendency to use red and green for structural purposes or in contrast, but blue, yellow, brown and black will be used freely in such ways.

The subject's colour vision

He was tested with the following three tests, in order:

(1) *Ishihara Test.* He did the Ishihara Test, with the results given below, which showed errors on all of the twenty-four critical plates, and this test left little doubt that he was a protanope or protanomalous, because of the readings of '5' and '6' on purple dots to the right of the Plates *24* and *25*. This confirmed the impression given by his first comment that red had a different scale of values for him, because in his vision the red or left-hand figures are darkened so that they cannot be seen against the dark grey background dots.

Correct reading	Subject's reading or readings
12	12 in orange dots (strong) on pale green-grey background of dots
8	Faintly as '8' strongly as '3'. 8 in pink dots (poorly defined). Background dots yellow-brown, light red, etc.
6	Same colours but figure doubtful
29	70. Same colour scheme as above
57	35. Same colour scheme (poorly defined figures)
5	Indefinable
3	Appears to be figure '5'
15	Appears to be figure '17'
74	Figure '21' fairly plainly defined
2	Quite indefinable
6	Quite indefinable
97	Quite indefinable
45	Quite indefinable

Correct reading	Subject's reading or readings
5	Spots of yellows, dark reds and pinks
7	Spots of yellows, dark reds and pinks
16	Similar, with some green spots
73	As the last plate
None	Possibly a figure '5'
None	Figure '2' rather better defined than no. 18
None	Might be '45'
None	Uncertain
26	Red spots in left half, purple spots on right half of circle. Do not form recognizable pattern
42	Similar to the last but different arrangement
35	Figure '5' in purple spots at right
96	Very undefined. Might be '6' in pink and purple at right

(2) *The Farnsworth Dichotomous Test* (Panel D-15). In this the subject has to set fifteen coloured caps in order, starting with a blue test cap, *P*. His order was as follows: *P, 1, 14, 15, 2, 3, 4, 5, 6, 9, 8, 13, 12, 11, 10, 7.*

This is not a typical arrangement of errors for any kind of colour defective, but is sufficient for absolute failure, because caps *14* and *15* are displaced from the end almost to the beginning, while a series, *2* to *6*, are displaced from the beginning towards the middle, and there are several other transpositions. In the normal or correct order, the caps pass from blues through blue-green, green and yellow-green to yellow, reddish and purple colours. The displacements he made imply that he confused purples with blues, and green with reddish and yellowish hues.

(3) *Pickford–Nicolson anomaloscope.* In the red–green test it was possible for this artist to match red or green, or any mixture of these coloured lights, with the yellow, provided that the yellow was sufficiently darkened for the pure red match, nearly to black, and stepped up in brightness steadily through orange and yellow to green. If the brightness were carefully matched for him, there was no point at which the red or green or the mixture of them seemed different from the yellow. Thus he is a dichromat with shortening of the red end of the spectrum—a protanope, which confirms the Ishihara Test.

In the blue–green test he did not differ from the normal, and in the yellow–blue test he differed only in requiring slightly more blue than the normal in the mixture of yellow and blue. This might be due to reduction in his sensitivity to blue as a result of yellow pigmentation in his optical system owing to his age being over sixty.

He had noticed in his own colour judgements in art and everyday life that there was a certain tendency to fluctuations in his colour vision. He said he had always suffered from visual fatigue when he had to do close work. In the anomaloscope test for red and green he felt that there might be some colour differences if he rested his eyes, but he was never certain that they were present when he looked carefully at the test spot.

The artist was not tested with the Nagel anomaloscope, because he lived about 400 miles from the Psychology Laboratory, and this instrument could not be carried to his house.

The artist was kind enough to answer a number of questions about his colour vision and his art and they are given below, with his answers.

Answers by a protanopic artist to questions about his art

You went to a school of art?	Yes.
Did you find you had unusual colour vision there?	No.
When did you find out that your colour sense was unusual?	Gradually. When I was in a commercial studio in Sydney, about 28 years of age I used to get a bit confused, and I noticed that small areas of colour gave me trouble. When I was playing snooker the brown and green balls were confusing; not the red or black.
What colours gave you most trouble?	Orange and green on the one hand, and lavender and blue on the other.
Red has a different scale of values for you, I believe?	In the distance it does not stand out anything like so much as violet or blue.
Did the knowledge of your unusual colour sense affect your art?	I don't think so because I had no means of judging what other people saw.
Does it affect your ideas about other people's art?	I find myself agreeing with the colour choices of other artists. It is all so individual; we have our likes and dislikes and there is no more to be said.
What do other people think of your paintings from the point of view of colour?	Some like them and some don't. I exhibited two of those you photographed and people seemed to like them.
You are not aware of compensating for any deficiency?	It is difficult to say. I look on myself as the sole judge. If I can please myself I'm satisfied. But I never can do that fully. For fifteen years I've done nothing but black and white work. It happened, owing to circumstances.
Black and white art was not adopted because of your colour vision being unusual?	No! I was offered work after the war in maps, and illustrations.
You never earned your living as a colour artist?	Oh! Yes! I did! I'll show you some of the hideous tripe one does to keep the pot boiling, if you like.
Yes! Please do! (Here he showed several portfolios of calendars and Christmas cards.)	Two Christmas cards and one calendar were best sellers; Dickensian sort of stuff. They have lots of colour in them.
You notice you don't use reds much?	Oh! I wouldn't say that. I wouldn't agree with you. There is a lot of orange-red, and there are brightly lighted windows. Many of the cards are in black and red, for instance.
But that is simply the printer's ink.	Yes! But, if you want your red to tell out you put it on a small area. This is more an artist's dodge than something due to my colour vision.
Is your colour vision variable from time to time?	What I see often varies considerably. If I look at a thing and say it's orange or lime green, and then look away and back again, it may seem different. I suppose these changes occur from day to day. There's one thing: I've been absolutely honest with you, but it's hard to give a decisive answer sometimes to some of your questions.

In summing up, it could be said that nothing unusual would be noted about this artist's colour work unless we knew that he was a protanope. He is extremely skilful in using colours, and never makes a blunder as many colour-blind people do. Apparently he is not conscious of any kind of deliberate adjustment or compensation for a defect. It would be noticed that he uses browns, black, ochres, yellows, oranges and greens, where they are appropriate, and to give highly sensitive colour harmonies and contrasts. One might notice that he avoided red, supposedly because of aesthetic taste, although he himself would not think that he avoided either red or red–green contrasts. When it is understood that he is a protanope, it is clear that his use of colour falls into line with the special characteristics and limitations of his colour vision.

If we ask what are the characteristics of his colour vision, on the most widely accepted views at the present day, he is like the chemist Dalton, and lacks sensitivity to red altogether. His vision must be made up out of varying saturations and combinations of what we should call green and blue. This interpretation is based on the view that all protanopes absolutely lack the red-responding substance in their eyes. For this artist, however, the red end of the spectrum appears to be less completely darkened than it is for some protanopes, although he can match any red with yellow or green if appropriately darkened. This would suggest that all protanopes are not quite alike, and that in his case the red-responding substance may be present to a small extent, but that it excites sensations the same as the green-responding substance, owing to a confusion of neural mechanisms or pathways. This is a 'collapse' form of colour vision defect, and its presence as part of his abnormality may accord with the fluctuating tendency in his red–green vision, which has been noticed in a number of colour-blind subjects. Sometimes the additional sensitivity comes into action, but it is readily put out of action by fluctuations of function and fatigue. Thus he is more responsive to the yellow and red parts of the spectrum than many protanopes, but this does not mean that he sees these parts as different from the green parts, except that they are darker for him. A painting in which he used red freely in colouring women's dresses was a copy of another painting from a black and white half-tone reproduction, in which he made a harmony of warm colours.

Note on the Colour Plate

The colour plate, part of the cost of the blocks for which was met by the Carnegie Trust for the Universities of Scotland, gives one painting by each artist. No. 1, 'An Indian Gentleman', by Mr Ray, shows his use of red, blue, black and yellow, with more green than is usual in his work. No. 2, 'Puerto de Soller Majorca', by the protanope, illustrates his characteristic range of colours, in this case very true to nature. Blues appear for sea, sky and distant mountains, while brown, grey and fawn are used on the near hill-sides. There is almost no red, and only a suggestion of grey-green for the cork and olive trees.

Discussion

In the previous paper the influence of deuteranomalous colour vision on a painter's art was explained. His preference in painting was to use subdued colours in which red and green were used very little or not at all, but he made many paintings with strong

1

2

constructional patterns in yellows and blues because he found that people liked them. In some paintings he introduced strong red colours, and in others rather less marked greens. Coloured illustrations of some of these paintings were given.

In this paper it is possible to see the influence of protan defects of colour vision on the work of two artists. The first, a protanomalous subject, uses red, gold and blue to build strong designs, but very seldom makes any use of green. It is certain that the reds he uses must be much darkened in comparison with the way we see them, and yet his paintings are aesthetically satisfying to us. The absence of greens appears as if it were a matter of aesthetic taste, and yet it must be determined at least in part by the relative weakness of red–green contrasts for him. Red and green will be to a certain extent like yellow for him, and he has chosen to stress most the one, namely red, which is for him the less visible of the two. Nevertheless he likes greens. Yellows look much more like greens to him than like reds, and this may be why he can employ them to provide contrasts with reds.

In the protanope's art we see an extremely skilful avoidance of possible difficulties and faults which might be caused by his peculiarities of colour vision. Red, as for Mr Ray, is greatly darkened for him, and red or green, in so far as red can be seen at all, are indistinguishable. What he calls yellow, orange and red must be different intensities and saturations of green, except that there is evidence that he is not an absolute 'loss' type of protanope, but may see red light to some extent—but it will not be distinguishable from green except in brightness or saturation. Unlike Mr Ray, he seldom, or very rarely, uses what we should call saturated reds, but substitutes colours which to us are dark oranges and yellows. Greens he does use. One can understand a red–green defective artist avoiding the use of either red or green, because they are not as clearly different as to us (for a deuteranomalous artist like Mr Purdy or the protanomalous artist, Mr Ray), or they are completely alike (for the protanope, if equated for brightness). The three artists considered have solved the problem each in his own way. By preference, Mr Purdy uses neither red nor green, but he does sometimes employ the one and sometimes the other, without using them together. Mr Ray seldom uses green. The protanopic artist almost never uses red. All their work is a tribute to the skill with which they can avoid the difficulties. It is something like the language of a person who carefully avoids using words of which he is not sure of the meanings. It appears to the writer that this skill has been attained unconsciously by Mr Ray and the protanope artist, but rather more consciously by Mr Purdy, who had some training in psychology and had clearer conscious appreciation of the peculiarities of his colour vision. Mr Ray also had some training in psychology. Absolute mistakes of colouring, such as giving a man a green face, or making the petals of a rose sky-blue, would not be artistic achievements unless they were intentional and in some way aesthetically appropriate. Intentional and aesthetically appropriate changes of colour of this kind, such as making a horse blue, however, would be done better by an artist who was not colour blind, and therefore, as seen in this paper, the peculiarities of colouring used by colour defective artists tend to be due to skilful avoidance of errors and adaptations of their own colour choices to be satisfactory also for normal colour vision, rather than the exploitation of uncontrolled blunders.

It is interesting that the protanopic artist wrote in answer to a question about the

artist Léger (1881–1955); 'I have known the work of Fernand Léger a long time and always admired it very much.' Léger may have been the colour blind artist, 'C.E.', of Strebel (1933)). He was possibly a protanope. Presumably some harsh contrasts in Léger's work, such as those of the red and black bands or stripes in a painting called 'Yellow Flowers in a Blue Vase' (1950), would be less harsh and perhaps more satisfactory in an aesthetic way to a protan than to a person with normal colour vision. Approximately what the protan sees may be observed by viewing such a painting through a blue-green (minus-red) filter of glass or gelatine. Then it is like other paintings by Léger, such as 'Leaves and Shell' (1927).

An Extreme Protanomalous Artist

Since this article was written another protanomalous artist has been found. He is a fifteen year old school boy who has unusual artistic ability. He knew he had defective colour vision because of difficulty with red, brown and green billiard balls at snooker. He was tested in the Psychological Laboratory by Miss Mollie G. Mackinnon, and carried out the following tests: (1) The Ishihara Test (result – 21 errors on 24 plates; protan); (2) The Nagel Anomaloscope (red-green test result– 49 ± 23; normal reading 40; (3) The Pickford-Nicholson anomaloscope (red-green test result – $20{\cdot}5 \pm 20{\cdot}5$; normal reading 25). These tests showed that he was an extreme protanomalous subject. He had a very large matching range which included the normal mid-matching point, and the red end of the spectrum was shortened for him in a characteristic way. His paintings were usually in browns, blues and yellows, with occasional patches of red or yellowish green. They were much more like those of the protanopic than of the simple protanomalous artist, and this is what would be predicted, provided he did not make outright blunders, because he has very poor red-greed hue discrimination. It is interesting that a boy of his age had already learned to avoid outright blunders of colour in his art.

References

Collins, M. (1925). *Colour Blindness.* London: Kegan Paul.

Kalmus, H. (1965). *Diagnosis and Genetics of Defective Colour Vision.* London: Pergamon (in the Press).

Pickford, R. W. (1951). *Individual Differences in Colour Vision.* London: Routledge and Kegan Paul.

Pickford, R. W. & Lakowski, R. (1961). The Pickford–Nicolson anomaloscope for testing and measuring colour sensitivity and colour blindness and other tests and experiments. *Brit. J. physiol. Optics*, **17**, 131–50.

Pickford, R. W. (1964). A deuteranomalous artist. *Brit. J. Psychol.* **55**, 469–74.

Pickford, R. W. (1965). The influence of colour vision defects on painting. *Brit. J. Aesthet.* **5**, 211–26.

Riddell, W. J. B. (1949). Discussion on colour vision in industry. *Proc. R. Soc. Med.* (*Ophthalmol.*), **42**, 145–50.

Strebel, J. (1933). Prolegomena optica zum Bildnerischen Kuntschaffen. *Klinische Monatsblatter fur Augenheilkunde*, **91**, 258–72.

Trevor-Roper, P. D. (1959). The influence of eye disease on pictorial art. *Proc. R. Soc. Med.* (*Ophthalmol.*), **52**, 721–44.

(*Manuscript received 24 November 1964*)

Brit. J. Psychol (1965), 56, 4, *pp.* 431–437
Printed in Great Britain

SKIN CONDUCTANCE CHANGE AND SENSORY DISCRIMINATION

By R. S. CORTEEN and A. R. BLACKMAN*
Psychology Department, University of Edinburgh

The relation between mean log. change in conductance and sensory discrimination was investigated. Highly significant relations were found with CFF and two-point tactile thresholds. No significant relation was found with discrimination of a pure tone from background white noise. An attempt at explanation is offered in terms of a cortically controlled centre mediating a sensitizing or orienting response of which conductance change is a peripheral manifestation.

Apart altogether from specific sensory defects, individuals vary in their ability to make sensory discriminations. There seems to be some evidence that this inter- and intra-individual variability is a function of a common central process mediating all sensory reception (Duffy, 1962; Venables, 1962). Covian, Timo-Iaria & Marseillan (1961) state that 'activity in sensory systems is subject to modification by central regulatory mechanisms and background states, sufficiently pronounced to be capable of influencing perception and so to be of interest for psychology'. Recent discoveries concerning the physiology of the nervous system, which will be briefly reviewed later, seem to indicate that this process is located in the non-specific reticulothalamic system. It therefore seems reasonable to assume that if we can measure some manifestation of relevant activity in this system we should find that it relates to sensory discrimination.

A recent review by Duffy (1962) and some reported research by Venables (1962) seem to imply that sensory discrimination is related to general activation level. The evidence reviewed by Duffy is mainly indirect but on the basis of it she hypothesizes a curvilinear relation between activation and sensory sensitivity. Venables related resting skin potential to temporal visual discrimination and reported correlations of between $-0{\cdot}64$ and $-0{\cdot}85$, but only for subjects falling within a certain range of negative potentials. He found no evidence of a curvilinear relation.

Resting skin potential has sometimes been observed to be related to resting conductance levels, but there is also evidence that no such relation exists (Tolles & Carberry, 1960). This is emphasized by the fact that Eysenck & Warwick (1964) found no relation between basal skin conductance level and two-flash thresholds. The experiments to be reported here, however, were initiated on the assumption that a relation would be observed between basal conductance level, generally assumed to be a measure of general activation level (Malmo, 1959), and sensory discrimination. In all cases measurement of resting conductance level was made with apparatus using an a.c. source, similar to but more sensitive than the apparatus used by Tolles & Carberry (1960), in order to eliminate problems of electrode polarization.

In a pilot study, using thirty-three male subjects between the ages of 15 and 18 yr, no relation was observed between two-point tactile thresholds, measured on the nape of the neck, and basal skin conductance level. When, however, the mean skin

* Now at Department of Psychology, McGill University, Montreal, Canada.

conductance change (GSR) in response to twenty words read to the subject was calculated, this measure was found to be significantly related to the two-point threshold. The difference between the mean thresholds of the group with above-average GSR level and the group with below-average GSR level was highly significant ($t = 3{\cdot}5$, D.F. 31, $P < 0{\cdot}01$).

This finding was somewhat unexpected, particularly in view of Venables' findings and conclusion. An obvious assumption was that GSR is a manifestation of the central mediating process mentioned above. But the experiment was not decisive enough to have any such far-reaching conclusion based upon it, and it was decided to go ahead with a more thorough investigation utilizing a more sophisticated determination of sensory thresholds.

The experiment reported here was therefore designed to investigate relations between (*a*) basal conductance level and visual, auditory, and tactile discrimination thresholds, and (*b*) mean GSR (log. change in conductance) to twenty words and visual, auditory, and tactile discrimination thresholds.

Method

Subjects

Sixteen male undergraduates from the Psychology Department of the University were used as subjects. Their ages ranged from 19 to 32 (average 22·5 yr). None had any known visual or auditory defects.

Apparatus

The apparatus for measuring conductance level and change consisted basically of a Wheatstone bridge unit fed by an oscillator. The current from the bridge was amplified before passing into a quick-response pen recorder. Two arms of the bridge were fixed resistances; the third arm was a variable resistance and capacitor in parallel, adjustable to match the impedance of the subject; the fourth arm was the palmar resistance of the subject which in fact, with alternating current, showed the equivalent circuit characteristics of a resistance in parallel with a capacitor. The variable resistance was judged to be equal to that of the subject when the width of the oscillatory trace on the recording paper was at a minimum (theoretically zero, but practically a few millimetres at maximum sensitivity). The GSR were shown as characteristic changes in the width of the oscillatory trace. At maximum sensitivity the apparatus was capable of discriminating a change of 20 Ω on a basal resistance of 20,000 Ω.

Two silver–silver chloride electrodes were screwed into the arm-rest of the experimental chair, so that the subject could comfortably rest his right hand, palm downwards, on them. The electrodes were saturated with 3 % sodium chloride solution before each recording, and were kept in clean water between experimental sessions.

The photic stimulator used to measure the critical flicker frequency (CFF) consisted of a xenon-filled stroboscope lamp seated in a reflector covered by a perspex diffusing screen, and a control unit. The perspex screen was covered by a sheet of dull black paper with a centrally placed, $\frac{3}{4}$ in. diameter aperture. The position of the stroboscope unit was adjusted so that the aperture was at the subject's eye-level, and at a distance of approximately 36 in. from the subject. The flash length was a constant 65 μsec; the flash-rate was variable over the range 2–60 c/s, and the intensity of the fused light was 180 ft.-lamberts. It should be noted that whilst the flash length was constant, the light/dark ratio varied directly with the flash-rate.

A 'Peters' basic diagnostic audiometer was used for measurement of the auditory discrimination threshold. This instrument was set to generate a pure 1000 c/s tone which could be passed into either headphone, the intensity of the note being variable. An auxiliary unit also manufactured by 'Peters' was used to generate white noise which was passed into both earphones. The audiometer is so calibrated that it measures hearing loss, not the actual intensity of the note passed through the earphones. The pre-set threshold intensities (without the white noise)

are those just perceptible to a test group of subjects between the ages of 18 and 25 years with healthy ears when listening in completely silent surroundings. The white noise provided a constant background from which the single note had to be discriminated, and it served to mask the many extraneous noises which are inevitable outside a sound-proofed room.

The device used to obtain the two-point tactile threshold was a simple aesthesiometer: a pair of compasses with the needles replaced by blunted ebonite pins at right-angles to the arms. A metal strip, calibrated to give the separation of the points in millimetres, was attached to one of the arms. The handle of the aesthesiometer was lightly clamped so as to allow freedom of movement only in the vertical plane. This ensured approximately equal pressure of the points on the skin over all the trials.

Procedure

The subject, after washing his hands, was seated in the experimental chair, and his right hand placed on the arm-rest so that the two electrodes were in contact with the palm. A strap was tied over the hand in order to minimize movement. The subject was asked to sit comfortably and relax while the experimenter adjusted the controls of the recording equipment. Light conversation was kept up, but questions as to the nature of the experiment were not answered, except to reassure subjects that no noxious stimuli were to be used.

The following instructions were then given to the subject: 'I am going to read out a list of ordinary words to you. They will come at about 20–30 sec intervals. You should *not* say anything in response to the words, just think about them. Please try and avoid moving your right hand. Any questions? . . . Please close your eyes, now, and relax; do not speak, just think about the words as you hear them.'

As soon as the subject's level of resistance had stabilized the first word was read out. The subject's response (the drop in resistance) was recorded, and the next word read out when the level of resistance had again stabilized. This procedure was repeated for each of the twenty words in the list (list 1). The sensitivity of the preamplifier was adjusted to maximize the trace-width during the GSR. The variable resistance and capacitance components of the bridge were set before each word to give a minimum trace width.

The recorder was equipped with an event-marker which was activated by the experimenter at the moment of stimulation.

Exactly the same procedure was carried out at the end of the experimental session, using a different list of words (list 2). In the intervening period, the subject's right hand was released from the strap, partly to prevent discomfort and partly to avoid an increase in the temperature of the palmar surface.

The aim was to use neutral words as stimuli; those chosen were considered to have a minimal emotional connotation.

In the interval between the two sets of recordings, measurements were made of the subject's discrimination thresholds in three sensory modes.

(*a*) *Visual*. The critical flicker frequency (CFF) was chosen as the measure of visual sensitivity. The subject's task was to tell the experimenter, who was slowly increasing the frequency of flicker of the stroboscope lamp, when he thought the flickering light had fused. A similar procedure was carried out while the frequency was slowly decreased from well above the CFF; the subject was asked to indicate when he could first distinguish regular flickering. Five trials were given in each direction, the first in each case being considered a practice trial and the result discarded. The subject's CFF was taken as the average of the eight remaining trials.

(*b*) *Auditory*. The subject's task was to discriminate a 1000 c/s tone from a background of white noise. A constant level of white noise was fed into both earphones and the single tone into either the right or left earphone. Again the method of limits was used: the intensity of the single tone was decreased until the subject indicated that he could no longer hear it; it was then gradually increased from a subthreshold value until the subject reported hearing it again. Four trials were given in each direction and for each ear, the results from the first trial in each case being discarded. The measure of auditory sensitivity used was the average threshold value over the remaining twelve trials. The units of measurement were to some extent arbitrary, as they relate to stands derived from measurement of absolute threshold, but this was immaterial since a relative rather than an absolute measure was required.

(*c*) *Tactile*. The measure of sensitivity in this mode was the two-point threshold. The area of

skin chosen was that at about 3 in. above the wrist on the inner aspect of the left fore-arm. The following instructions were given to the subject: 'Please close your eyes. I will be placing either one or two points on your skin just above the left wrist. Tell me in each case whether you think it is one or two points.' Fifteen practice trials were given at different distances of separation between 0 and 60 cm. Then twenty-five trials were given at 1 cm intervals centred around the threshold as estimated from the practice trials. The distances were given in a randomized order, and the two-point threshold was taken as that distance of separation which the subject judged to be 'one point' on 50% of the occasions.

The experimental room was kept at a temperature between 20° and 22° C. The experimental session lasted approximately 3–4 hr.

Results

Table 1 shows the 'mean GSR' (variable 1) for each of the sixteen subjects expressed as the logarithm of the average change in conductance to the stimulus words of list 1 and list 2. Variables 2–4 are the corresponding average threshold values. There was no marked heteroscedasticity in the distributions, and product-moment correlations, also shown in Table 1, were calculated between the threshold variables (2, 3 and 4) and mean GSR (1), after checking that there were no significant quadratic relations.

The curves of best fit were:

Variables y	Variables x	Regression equations
2	1	$y = 1{\cdot}748x + 49{\cdot}323$ ($y = -2{\cdot}190x^2 - 0{\cdot}782x + 49{\cdot}013$; significant linear relation)
3	1	No significant linear relation
4	1	$y = -10{\cdot}043x + 25{\cdot}308$ ($y = -7{\cdot}552x^2 - 18{\cdot}770x + 24{\cdot}236$; significant linear relation).

There was no significant correlation between any of the measures of sensory discrimination.

Discussion

The results obtained indicate a relation between sensory discrimination and mean log. change in conductance (mean GSR) only when the discrimination task is one which requires the subject to discriminate between two or more spatially or temporally proximate stimuli. This fact seems to rule out explanation in terms of a common intervening psychological variable such as confidence level or willingness to risk a judgement. Explanation in these terms would fail to explain the non-significant relation between mean GSR and the measure of auditory discrimination.

It would be equally misleading to attempt to explain the results in terms of general activation level. Malmo (1959) and others have pointed out the relation of basal conductance level to activation but there has been no observed linear relation between mean log. conductance change and basal log. conductance level.

It might be more reasonable to assume that the GSR is a peripheral manifestation of a sensitizing or orienting response (Berlyne, 1960) which occurs more or less independently of the general level of activation; though Oswald's (1962) observation that GSR tend to disappear at the onset of sleep, only to reappear later, does indicate some non-linear relation between the orienting response and general activation level.

The idea that such an orienting response might be related to sensory sensitivity is supported by several findings. Edelberg (1961) found that tactile sensitivity is enhanced immediately following a GSR and Lowenstein (1956) demonstrated that sympathetic afferent activity increased the sensitivity of tactile receptors. These two findings led Martin & Edelberg (1963) to hypothesize that the GSR may be part of a screening mechanism by which the organism can regulate its sensitivity in accordance with its tendency to reject or accept stimuli.

Table 1. *Mean GSR and average values of measures of visual discrimination (CFF), auditory discrimination* (1000 c/s *tone from white noise*) *and tactile discrimination* (*two-point threshold*) *for sixteen subjects.*

	Variable			
Subject	(1) Mean GSR (log. change in conductance)	(2) Visual (c/s)	(3) Auditory (db)	(4) Tactile (cm)
1	−1·198	46·88	40·88	35·5
2	−0·490	49·38	38·56	37·5
3	−1·046	46·00	36·82	38·0
4	0·223	47·75	40·50	15·0
5	−0·564	49·17	37·81	23·0
6	−0·563	47·88	31·56	34·5
7	−0·079	49·38	34·50	31·0
8	−0·198	50·33	36·50	32·5
9	−0·798	49·25	37·83	34·5
10	−0·643	47·50	35·38	26·5
11	−0·097	49·13	38·56	29·5
12	−1·331	47·13	38·00	38·0
13	−0·049	49·50	34·88	20·5
14	−0·301	49·38	34·87	24·5
15	−0·946	47·38	35·63	32·5
16	−0·292	48·50	35·07	36·0
Mean	−0·52	48·41	36·71	30·56
S.D.	0·43	1·17	2·33	6·64
r_{1y}	—	+0·650	−0·153	−0·657
P*	—	< 0·01	> 0·05	< 0·01

* For differences of the correlation coefficients, r_{1y}, from zero.

Jung (1958) found that repeated thalamic and reticular stimulation increased the CFF of specific neurons by 10–15 c/s. The same author also found that stimulation of the non-specific thalamic system led to an increase in the amount of neuronal discharges and the number of neurons responding to retinal afferents in the visual cortex. These findings by Jung indicate that the central structure largely involved in this sensitizing or orienting response is located in the rostral part of the non-specific reticulo-thalamic system, that part which, according to Jasper (1958), 'seems to mediate the rapid, short-lasting, or phasic activation of the cortex'. That this area is also involved in facilitation of the GSR is confirmed by Wang (1957, 1958).

It is likely that the cortex plays a selective inhibitory role upon the strength of the orienting response. This is indicated by the evidence of Wang, Stein & Brown (1956) and Wang (1958) who noted that the GSR of a cat is enhanced after removal of the forebrain, and by Jouvet (1961) who found that habituation of the orienting response

J. Brit. Psychol. (1965), 56, 4, pp. 439–446
Printed in Great Britain

SEQUENTIAL PART PRESENTATION: A METHOD OF STUDYING VISUAL FORM PERCEPTION

By JOSEPH H. McFARLAND*
Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts

Everyday perception of a form's parts as simultaneous and joined is viewed as a perceptual achievement dependent on the central operations of analysis and integration. Experiments have been conducted which employ a method of stimulus presentation which assures that neural response simulates the hypothesized analysis operation, viz. sequential part presentation. Perceptual measures are employed which are assumed to reflect the hypothesized integration. In one experiment, interpart intervals between the sequentially presented sides of a line triangle were varied and 50 % points determined for judgements of simultaneity and joining between the lines. In a second experiment, these thresholds were also determined, but for side part versus angle part presentation of the line triangle. With side part presentation, thresholds for both judgements of simultaneity and joining were found to be higher. It is concluded that normal perception of form, i.e., perception of parts as simultaneous and joined, can take place at larger intervals when sequential response is to side parts.

In studying visual form perception, it is customary to present the parts of a stimulus form simultaneously. Although this method is in common laboratory use, a theoretical rationale for the method is seldom, if ever, provided in current theorectical treatments of form perception.

While searching for such a rationale, it has become clear that the Gestalt treatment of form perception, though subject to diverse criticisms (cf. Hebb, 1949; Helson, 1926; Lashley, Chow & Semmes, 1951; Nagel, 1963; Piaget, 1952) does provide an argument for employing this method. In brief, the argument can be developed as follows. Forms in our everyday experience appear with their parts simultaneous in time and joined in space. Thus, in terms of the principle of isomorphism, the central nervous system processes simultaneously all portions of retinal stimulation arising from a form (cf. Koffka, 1935). Accordingly, in studying visual form perception, it is appropriate to present all parts of the form simultaneously.

If this is the rationale tacitly employed by many investigators, then a paradoxical situation has indeed been created. While in theory progress has been made beyond the Gestalt heritage (cf. Gibson, 1950; Hebb, 1949; Lashley, 1951; Piaget, 1963; Uhr, 1963), a Gestalt rationale is still employed in choosing a method of stimulus presentation for the study of form perception.

In view of this situation, two objectives are indicated: (1) the development of a rationale, in terms of present theory, for the method of simultaneous part presentation; and (2) the development of new theoretical models and theoretically appropriate methods of stimulus presentation.

In pursuance of the latter objective, three steps have been taken which are reported here. First, an approach has been developed which treats form perception as dependent on central operations of analysis and integration. Secondly, a method of stimulus presentation has been devised which is appropriate to this approach. Finally, experiments have been conducted which indicate the heuristic value of this

* Now at University of Denver, Colorado, U.S.A.

approach and method for providing information on the process of visual form perception.

As to the approach, visual form perception is conceptualized as dependent on a central mechanism that performs operations of analysis and integration. Analysis entails the production of sequences of responses to line and to angle portions of retinal stimulation; integration entails the production of a response to the products of analysis.

In terms of this theoretical model, everyday perception of a form's parts as simultaneous in time and joined in space is considered the perceptual achievement of these operations of analysis and integration. While a form's parts may be simultaneous and joined in terms of retinal stimulation, the perception of simultaneity and joining is viewed as dependent on the postulated central mechanism which produces a sequence of responses to parts—analysis; and which then produces a unitary response to the sequence of responses—integration.

The method of stimulus presentation that has been devised consists of presenting the parts of a form sequentially. This is considered an appropriate method for two reasons. First, it can be assumed that the response of the visual system to a form presented by this method simulates, in part, the hypothesized analysis entailed in everyday form perception when the form is presented with its parts simultaneous. Secondly, it is assumed that perception of the stimulus form presented by this method is the achievement of the hypothesized operation of integration. Thus, with this method of stimulus presentation, appropriate variation of stimulus parameters permits variation of the analysis operation and the consequences for integration, as manifested in perception, can be studied.

Two experiments employing this method are reported. In the first experiment, the intervals between the sequentially presented sides of a form are varied from 0 to 300 msec and thresholds for non-simultaneity and non-joining of parts are determined. These threshold intervals are assumed to define the limiting intervals at which normal form perception can occur for a sequence of responses to side parts.

In the second experiment, two conditions are employed. In one condition, as in the first experiment, a form's side parts are presented sequentially; in the second condition, the same form's angle parts are presented sequentially. Here also the intervals between the sequentially presented parts of the form are varied. The aim of this experiment is to provide information on the limiting intervals at which normal form perception can take place for a sequence of responses to side versus angle parts.

EXPERIMENT I

Method

Stimulus

The stimulus was a luminous (0·44 ft.-lamberts) equilateral line triangle. Sides were 1 degree in length by 4 min. in width at 17 in. The form was presented on a dark background, straight ahead at eye level, with one side in a horizontal orientation. Sides were flashed in sequence for 10 msec each. The sequence employed was base, then right side, and then left side. Interpart intervals (1) between the base and right side, and (2) between the right and left sides, were maintained equal and varied from 0 to 300 msec in 25 msec steps. This method, it may be noted, is similar to one employed by Brown & Voth (1937) and, more recently, by Müller (1963).

Apparatus

The sides of the stimulus form were prepared as three 35 mm slides. These slides were independently projected on to a ground-glass rear projection screen by means of three Sylvania R1131C bulbs. (These bulbs were originally designed for use in photographic transmission and have fast rise and decay times.)

For each stimulus presentation there were five time periods to be controlled and/or varied: the duration of flash for each of the three bulbs and the two intervals between the sequential flashing of the bulbs. These time periods were controlled by five solid state electronic timers permitting variation from 1 to 1000 msec. (Reliability was $\pm 5\%$ and repeatability was $\pm 1\%$ over the entire range.) Flashing of the bulbs in the appropriate sequence with apppropriate intervals was achieved by an electronic logic circuit. (This equipment is now a standard item by the Polymetrics Division, Pennsylvania Optical Co., Reading, Pa.)

Design

The stimulus was presented by a method of limits with ten trials at each of the thirteen interpart intervals: five in an ascending, and five in a descending series. Thus, there were 130 trials for each subject. Four males and four females were employed in each of two series sequences, ascending-descending and descending-ascending, making a total of sixteen subjects. Two males and two females in each of the series sequences were retested 24 h later.

Procedure

All testing was carried out in a dark room and subjects were dark-adapted for 10 min prior to testing. During the dark adaptation period, subjects were instructed as follows: 'In a few minutes I will ask you to open your eyes and look straight ahead. I will give you a "ready" signal and then flash a line triangle on the screen in front of you. As soon as you have seen the form, close your eyes. When your eyes are closed, I want you to report on certain temporal and spatial relations between the parts of the form you have seen. First, tell me whether the sides are simultaneous, overlapping, or successive in time.—When the sides come on and go off together, they are simultaneous; when they come on or go off in succession, but at some time more than one is on, they are overlapping; when they come on one at a time and no two are on at the same time, they are successive.—Then tell me whether all the sides are or are not joined to make a perfect triangle. If some do not join, tell me which pair or pairs of sides do not join to make a perfect angle.—Always make this judgement as if the sides were simultaneous.'

After 10 min. of dark adaptation. testing commenced and was usually complete within 45–60 min. During the testing period, subjects' heads were supported by a chin rest and lateral head movements were partially restrained.

Results

The judgements of temporal relations show the typical functions for three category judgements (Fig. 1A). In contrast to the judgements of temporal relations however, the judgements of spatial relations show a complex, non-linear decline with increase in interpart intervals (Fig. 1B). These judgements initially decline, parallel with judgements of simultaneity, reach a plateau at 75–100 msec, and then slowly decline further. Examination of individual functions shows that this plateau is attributable in part to the fact that between 50 and 125 msec, twelve subjects show a paradoxical increase in judgements.

For the eight subjects retested 24 hr later, the overall forms of the functions are comparable for both test days. Product-moment correlation coefficients for the 50% points are $+0{\cdot}696$ for simultaneity, $+0{\cdot}328$ for succession, and $+0{\cdot}913$ for all ends joined. Only the correlations for simultaneity and all ends joined are significantly different from zero ($P < 0{\cdot}05$). While all day 2 thresholds are elevated, this difference is not significant at the 0·05 level as determined by variance analysis.

The limiting intervals for simultaneity and joining, viz. from onset of the first

until offset of the last part, were obtained by adding the three 10 msec part durations with the two interpart intervals at the respective thresholds, 41 and 52 msec. Thus, the mean limiting interval is 112 msec for simultaneous judgements which is significantly different from the limiting interval of 134 msec for judgements of all ends joined (Table 1). (Owing to the non-linear change in judgements for all ends joined, it was impossible to determine a 50% point for four of the sixteen subjects in the conventional manner. For these subjects, another measure was employed: the interpart interval at which judgements had declined to a minimum prior to showing the paradoxical increase referred to above.)

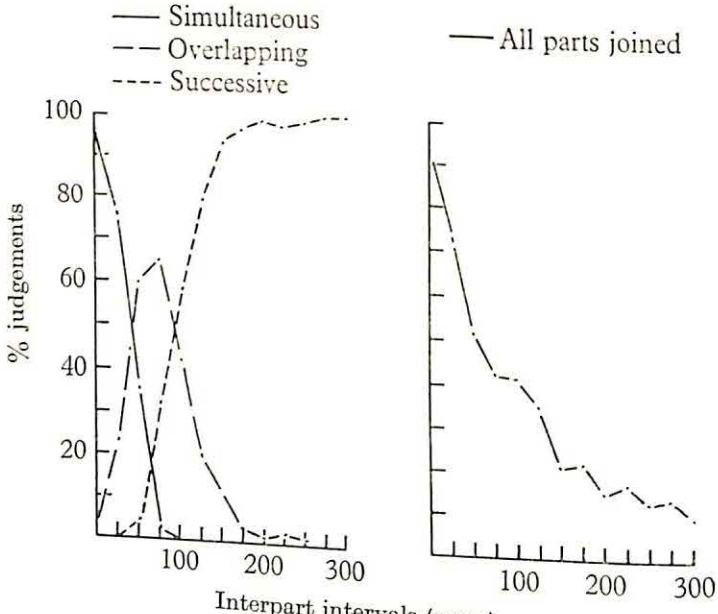


Fig. 1. Expt. I: mean percentage judgements of simultaneous, overlapping, and successive (A), and of sides joined to make a perfect triangle (B) as a function of interpart intervals (msec).

EXPERIMENT II

This experiment was basically the same as Expt. 1 with one main exception. Two conditions of presenting the triangle were employed: (1) sequential side part presentation, and (2) sequential angle part presentation. Accordingly, minor details of method, design, and procedure were changed.

For the angle part presentation, the three angles, formed by bisecting the sides of the triangle, were flashed sequentially, each for 10 msec. The sequence was right angle, then top angle, and finally, left angle. For both side and angle part presentation, the stimulus form was identical in all respects and the same as that employed in Expt. I. Two males and two females in each of the two series sequences employed in Expt. I were tested in each of two conditions sequences: side-angle part presentation and angle-side part presentation. Thus, there were sixteen subjects. Half of the subjects were tested at interpart intervals from 0 to 100 msec, with both conditions presented in the same testing session; the remaining half were tested at interpart intervals from 0 to 300 msec, with conditions separated by 24 hr.

Results

Examination of the data for the subjects who provided judgements for three categories of time relations at interpart intervals from 0 to 300 msec shows that all functions are basically similar to those obtained in Expt. I. Functions for simul-

taneity and all parts joined for subjects tested from 0 to 100 msec are similar to the corresponding functions obtained for subjects tested from 0 to 300 msec.

Comparison of the 'simultaneity' and 'parts joined' functions for the two conditions (Fig. 2A, B) shows that angle-part presentation significantly decreases both 50% points. It is also clear, from examining the limiting intervals (Table 1), that the direction of the difference between intervals for simultaneity and side parts joined is comparable to the difference obtained in Expt. I, but opposite to the direction of the difference between intervals for simultaneity and angle parts joined. This interaction is significant ($P < 0{\cdot}05$).

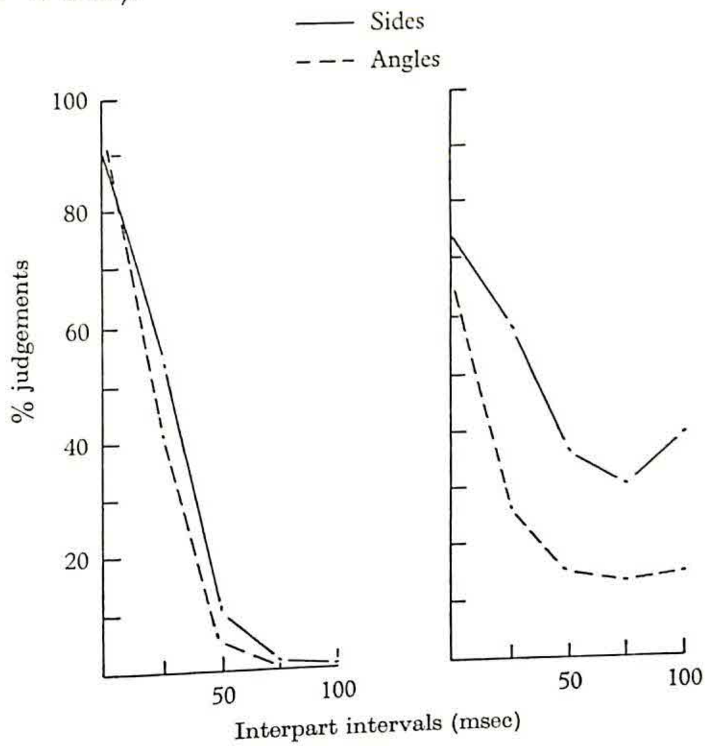


Fig. 2. Expt. II: mean percentage judgements of 'simultaneity' (A), and of 'parts joined' to make a perfect triangle (B), as a function of interpart intervals (msec) for side and angle part presentation.

Table 1. *Mean limiting intervals (msec) for judgements of simultaneity and of parts joined to make a perfect triangle, Expts. I and II*

(Limiting intervals, from onset of the first until offset of the last part, were obtained by adding the three 10 msec part durations to the two interpart intervals at the respective 50% points.)

Experiment	Parts	Simultaneous	All parts joined
1	Sides	112	134
2	Sides	84	102
2	Angles	71	52

Discussion

It will be recalled that everyday perception of form is viewed as the perceptual achievement of the central operations of analysis and integration. The operation of analysis produces sequences of responses and the operation of integration produces a

response to this sequence of responses. Normally, a form's parts are perceived as simultaneous and joined. The thresholds employed here are assumed to define the limiting intervals at which normal perception can be achieved to the hypothesized sequences of responses.

Viewed in these terms, the findings show: (*a*) as time between sequential responses increases, both perception of simultaneity and joining show an initial rapid and linear decrease; (*b*) when sequential response is to side parts, joining can be perceived over a larger interval than simultaneity; when sequential response is to angle parts, simultaneity can be perceived over a larger interval than joining; (*c*) when sequential response is to side parts, both joining and simultaneity can be perceived over larger intervals than when sequential response is to angle parts (see Table 1).

The fact that both perception of simultaneity and joining show a parallel decrease, as time between response increases, suggests that these two features of normal form perception depend on a common mechanism as proposed here. However, it is clear from the overall pattern of limiting intervals, that the operations of integration hypothesized to underlie these two features of form perception are not identical; they differ depending on whether responses to side or angle parts are integrated. Moreover, it is also evident that there is a differential integration underlying perception of line parts as simultaneous and joined versus perception of angle parts as simultaneous and joined.

Clearly the interpretation offered for these findings is subject to the reservation that, since eye movements are not monitored during stimulus presentation, eye movements between presentation of stimulus parts may be contributing to the definition of limiting intervals for simultaneity and joining. When, however, the limiting intervals for simultaneity (71, 84, 112 msec) and joining (52, 102, 134 msec) are compared with the smallest reported latency for a saccade (120 msec) (cf. Bartz, 1962; Ditchburn & Ginsborg, 1953; Ginsborg, 1953; Nachmias, 1959; Ratliff & Riggs, 1950; Tinker, 1958; Wertheimer, 1954), it is clear that saccadic eye movements are probable at only one (134 msec) of the limiting intervals. Furthermore, if saccadic eye movements were contributing to the perception of non-simultaneity and non-joining of all parts, one might expect to see a sharp decrease in responses of simultaneity and joining at or immediately after an interpart interval (45 msec) when the period from onset of the first until offset of the last part coincides with the period of latency for a saccade (120 msec.) When Figs. 1 and 2 are examined, however, this expectation is not supported. It may be added, however, that there is no evidence at present which excludes the possibility that eye movements contribute to the findings reported here via generation of peripheral and/or central extra-ocular feedback.

In order to facilitate comparison with the related findings of others, it may be noted that the findings reported here shed some light on two questions of current interest, viz. what are the characteristics of the neural mechanism determining (*a*) temporal relations, and (*b*) figural relations between visual stimuli?

With respect to temporal relations, it has been proposed that the period of the EEG alpha rhythm is a manifestation of the central mechanism determining apparent simultaneity between stimuli (McCulloch, 1951; Murphree, 1954; Stroud, 1955). The findings here show that only when response is to line parts in sequence does the limiting interval for simultaneity fall within the range of alpha periods; when response is to

angle parts the limiting interval for simultaneity is smaller than the smallest alpha period. This feature of the findings suggests that apparent simultaneity between stimuli is not a simple function of their occurrence within a centrally determined and fixed interval, but depends upon the mode of integrating different stimuli.

In regard to the problem of figural relations, Lashley's (1942) suggestion that a line is the simplest visual element is still pertinent. Recent electrophysiological work has indicated that certain cells in the visual cortex of cats respond optimally to line stimuli in different orientations (Hubel & Wiesel, 1959; Hubel, 1963). Thus, it may well be that form perception in humans depends on the output of such figural specific cells. Accordingly, one might speculate that perception of form entails a response to line parts of the form—analysis, and then a response to this response—integration (Hebb, 1963). The findings here might then be interpreted as supporting this speculation since perception of simultaneity and joining can be achieved at larger intervals when line parts are presented than when angle parts are presented. However, an alternative interpretation for these findings may be offered from the vantage point of the view proposed here. Simply that human form perception depends on both a differential response to line and angle parts of a form—analysis, and a differential integration of line and angle part responses.

At present the operations of analysis and integration described here are hypothetical. There is little support for their usage in terms of present knowledge of the human nervous system's functioning. Nevertheless, formulation of these hypothetical operations has led to a method and to experiments which have provided new information concerning visual form perception.

This work was supported at Clark University by National Institutes of Health, Grant R-6403 (A), U.S. Public Health Service.

References

Bartz, A. (1962). Eye-movement latency, duration, and response time as a function of angular displacement. *J. exp. Psychol.* **64**, 318–24.

Brown, J. F. & Voth, A. C. (1937). The path of seen movement as a function of the vector-field. *Amer. J. Psychol.* **49**, 543–63.

Ditchburn, R. W. & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature, Lond.*, **170**, 36–7.

Gibson, J. J. (1950). *The Perception of the Visual World*. Cambridge, Mass.: Riverside.

Ginsborg, B. L. (1953). Small involuntary movements of the eye. *Brit. J. Ophthal.* **37**, 746–54.

Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley.

Hebb, D. O. (1963). The semiautonomous process: its nature and nurture. *Amer. Psychologist*, **18**, 16–27.

Helson, H. (1926). The Psychology of Gestalt. *Amer. J. Psychol.* **37**, 25–62.

Hubel, D. H. (1963). Integrative processes in central visual pathways of the cat. *J. opt. Soc. Amer.* **53**, 58–66.

Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–91.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt Brace.

Lashley, K. S. (1942). The problem of cerebral organization. In *Biological Symposia*, vol. VII, ed. Kluver, H. Lancaster, Pa.: Catell.

Lashley, K. S. (1951). The problem of serial order in behavior. In *Cerebral Mechanism in Behavior*, ed. Jeffress, L. A. New York: Wiley.

Lashley, K. S., Chow, L. K. & Semmes, J. (1951). An examination of the electrical field theory of cerebral integration. *Psychol. Rev.* **58**, 123–36.

MCCULLOCH, W. S. (1951). Why the mind is in the head. In *Cerebral Mechanisms in Behavior*, ed. Jeffress, L. A. New York: Wiley.

MÜLLER, K. (1963). Der Aufbau figural-optischer Phänomene bei sukzessiver Reizung. *Psychol. Arb.*, no. 7.

MURPHREE, O. D. (1954). Maximum rates of form perception and the alpha rhythm: An investigation and test of current nerve net theory. *J. exp. Psychol.* **43**, 57–61.

NACHMIAS, J. (1959). Two dimensional motions of the retinal image during monocular fixation. *J. opt. Soc. Amer.* **49**, 901–8.

NAGEL, E. (1963). Wholes, sums, and organic unities. In *Parts and Wholes*, ed. Lerner, D. New York: Free Press of Glencoe.

PIAGET, J. (1952). *The Origins of Intelligence in Children*. New York: International Universities.

PIAGET, J. (1963). Le développement des perceptions en fonction de l'âge. In *Traité de Psychologie Expérimentale*, vol. VI, ed. Fraisse, P. and Piaget, J. Paris: Universitaires de France.

RATLIFF, F. & RIGGS, L. A. (1950). Involuntary motions of the eye during monocular fixation. *J. exp. Psychol.* **40**, 687–701.

STROUD, J. (1955). The fine structure of psychological time. In *Information Theory in Psychology*, ed. Quastler, H. New York: Free Press of Glencoe.

TINKER, M. A. (1958). Recent studies of eye movements in reading. *Psychol. Bull.* **55**, 215–31

UHR, L. (1963). 'Pattern Recognition' computers as models for form perception. *Psychol. Bull.* **60**, 40–73.

WERTHEIMER, G. (1954). Mechanisms of saccadic eye movements. *AMA. Arch. Ophthal.* **52**, 710–24.

(*Manuscript received* 16 *June* 1964)

Brit. J. Psychol. (1965), 56, 4, pp. 447–454
Printed in Great Britain

THE EFFECT OF STIMULUS CONFIGURATION ON THE INCIDENCE OF THE COMPLETION PHENOMENON

By ELIZABETH K. WARRINGTON
The National Hospital, Queen Square, London

The phenomenon of 'completion' of incomplete figures across hemianopic field defects was studied in relation to the nature of the stimulus forms used. While the highest incidence of complete responses occurred with simple geometrical forms a substantial number occurred also in response to other types of forms, including complex representational material characterized neither by symmetry, nor 'goodness' in the Gestalt sense, but depicting objects which are familiar and readily named. It was concluded that completion responses are most readily given to those forms with which the patient is familiar and which he would expect to see in complete form. Unfamiliar forms not readily named are infrequently completed.

Visual forms may be 'completed' across blind sections of the visual field by patients with cerebral hemisphere lesions. This was first demonstrated by Poppelreuter (1917), who presented geometrical forms tachistoscopically in such a way that half of each figure fell into a defective area of the visual field. Many patients nevertheless reported having seen a complete form. This completion of forms that could not in fact have been seen in their entirety was interpreted by Poppelreuter as a compensatory mechanism analogous to the normal process of completion of stimuli overlapping the blind spot (Helmholtz, 1924–25). Fuchs (1920) confirmed these observations and further claimed that the effect could be elicited only with simple, regular and symmetrical figures (a circle, ellipse, square, and star shape were all completed). Meaningful drawings (dog, face, bottle, butterfly) were never completed even though symmetrical. Letters, words, straight lines and homogeneous surfaces were likewise never completed. These findings, obtained on a single patient, were thought sufficient to controvert Poppelreuter's view of completion as imaginative filling in of the missing parts on the basis of previous experience. They led to an explanation in terms of Gestalt theory of perception, completion being the outcome of autonomous forces which tend to produce a symmetrical and well-balanced configuration. Koffka (1935) made use of Fuchs's work in his advocacy of Gestalt hypotheses.

Poppelreuter and Fuchs both report, without further comment, that some patients failed to complete incomplete forms. Warrington (1962) showed that failure to complete was common in hemianopic patients and established that the occurrence of completion was correlated with parietal lobe dysfunction and unawareness of the visual-field defect. Normal subjects and patients with cerebral lesions but no field defect behaved similarly to the non-completers. These results cast doubt upon the view of completion as normal and compensatory. Having established that certain patients show the phenomenon of completion it is relevant to determine whether all visual forms are completed and, if not, the characteristics which distinguish those that are completed from those that are not. The study that is reported was designed to elucidate the critical differences.

RESULTS

Completion of simple geometrical forms (Set A test figures)

The percentage distribution of completion of half-figures at threshold is shown for the hemianopic and non-hemianopic patient groups and the control group separately in Fig. 2. The percentage distribution of complete responses for the half-figures in the hemianopic group was bimodal: thirteen of the twenty-six patients

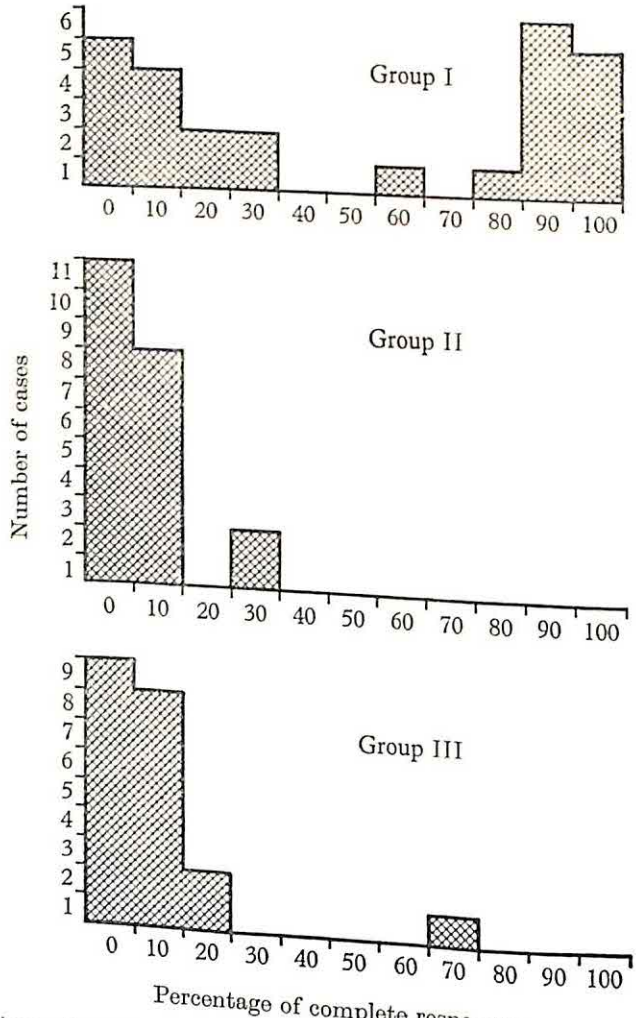


Fig. 2. Distributions of percentages of 'complete' responses by subjects in Groups I (with hemianopic visual-field defects), Group II (with damage involving cerebral hemispheres, but intact visual fields) and Group III (normal, subjects) to the half-figures of simple geometrical forms, set A, Fig. 1.

showed 60 % or more completion and thirteen showed 30 % or less completion. They may be regarded as two subgroups: (*a*) a completion group, i.e. those patients showing 60 % or more completion responses in the above distribution; and (*b*) a no-completion group, i.e. those thirteen cases in the above distribution showing 30 % or less completion responses.

The distribution of completion responses in Groups II and III (see Fig. 2) is regarded as comparable to that of those cases in Group I showing 'no completion'.

The mean percentage of completion responses for these two groups is 0 % and 10 % respectively and for the 'no completion' cases in Group I it is 10 %. The mean percentage and range of complete responses to familiar geometrical shapes is shown in Table 1.

Table 1. *Percentage and range of completion responses by Groups Ia, Ib, II and III to simple geometrical shapes, Set A*

	Group I(*a*) 'completers'	Group I(*b*) 'non-completers'	Group II	Group III
Mean % of completion	91	11	5	6
Range of completion responses (%)	60–100	0–30	0–30	0–70

Completion of other figures

(*a*) *Frequency of completion.* The percentage of completion responses by Group I (the hemianopic group) to figures 6–21 is shown as a frequency distribution in Fig. 3. The shaded circles indicate cases in which there was 60 % or more completion of simple geometrical shapes. Though the distribution here is less clearly bimodal, there is no overlap between those patients who were in the completion group and those in the

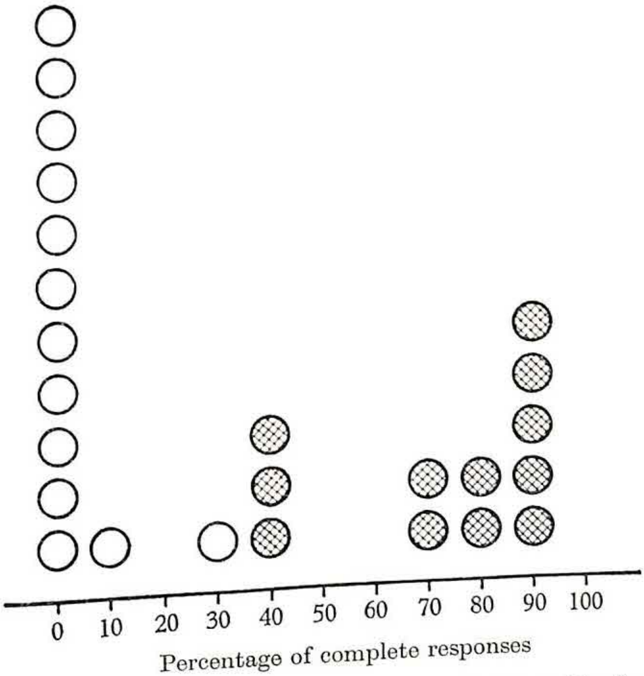


Fig. 3. Distribution of percentages of 'complete' responses in Group I to figures 6–21: ○, cases showing 30 % or less completion of familiar geometrical shapes; ⊕, cases showing 60 % or more.

no-completion group. Groups II and III showed no completion responses in this experiment. The mean percentage and range of completion responses for each group are shown in Table 2. For each subject-group the mean percentage of completion responses is lower than for the simple familiar geometrical shapes.

(*b*) *Distribution of completion responses to different type of test stimuli.* The mean percentage of completion responses given by Group I to the figures in Sets B, C, D and E (nos. 6–21) are shown in Table 3. The highest incidence of completion responses

was for Set D, which consists of detailed representations of objects. When the figures in Set C were completed the portion filled in was asymmetrical about the vertical axis, so that what was reported was the basic form in a rotated position (see Fig. 1).

Completion of animal figures

The animal figures (Set F, numbers 22–24) were shown to ten patients, six in Group I and four in Group II. Every subject in Group III was tested for completion with these figures. The only instances of completion responses were in the case of three patients in Group I. These three patients all showed marked completion of geometrical figures. The other three patients in Group I who did not complete the animal figures did not show any completion responses to geometrical figures.

Table 2. *Percentage and range of completion responses by Groups Ia, Ib, II and III to figures 6–21, Sets B, C, D and E*

	Group I(*a*) 'completers'	Group I(*b*) 'non-completers'	Group II	Group III
Mean % of completion	73	4	0	0
Range of completion responses (%)	40–90	0–30	0	0

Table 3. *Percentage of completion responses by Group I to Sets B, C, D and E*

	Set B (figures 6–9)	Set C (figures 10–13)	Set D (figures 14–17)	Set E (figures 18–21)
Group I	32 %	34 %	43 %	29 %

Table 4. *Ratio of 'naming' to 'descriptive' responses by Groups I, II and III*

Subgroups	Set B (nos. 6–9)	Set C (nos. 10–13)	Set D (nos. 14–17)	Set E (nos. 18–21)
Group I	0·48	0·80	0·89	0·77
Group II	0·57	0·87	0·95	0·86
Group III	0·47	0·85	0·96	0·87
All	0·50	0·84	0·94	0·83

Naming responses in Groups I, II and III

The incidence of responses incorporating a name for the figures in sets B, C, D and E are given separately and combined for Groups I, II and III in the form of a ratio of 'naming' to 'descriptive' responses. These values are shown in Table 4.

It can be seen that there is virtually no difference between the three subject groups in respect of the incidence of naming. For all three groups named figures amount to almost 100 % of the responses to Set D (nos. 18–21) but only between 40 % and 50 % of those in Set B (nos. 10–13). It can therefore be assumed that the figures in Sets B and D differ in the ease with which they can be named.

The relation between naming and completion

The number of completion responses for Group I as a whole was 319, and in 309 of these naming occurred. In eight cases only was the response limited to a description of the figures which were completed. They were all in Set B (nos. 6–9), the simple geometrical shapes lacking common names. The reverse, however, does not hold.

That is to say, naming occurred in a number of cases in which there was no completion. None the less, it was rare for completion to occur in the absence of naming. The figures in Set D, which were the most frequently named, were also the most frequently completed.

Discussion

Completion across hemianopic field defects of simple familiar forms has been compared with the completion of a variety of other visual forms ranging from unfamiliar geometrical shapes to drawings of familiar objects. Earlier findings that simple familiar geometrical shapes give rise to the highest incidence of completion responses were confirmed: 49 % of responses to the geometrical shapes were complete, while 36 % were complete for all the other test figures combined in Group I. Representations of objects with identifying detail produced the next highest incidence of complete responses, 43 %, while the simple but unusual geometrical shapes having no name in common usage and the complex shapes of objects without their identifying detail produced the lowest incidence of complete responses, 32 % and 29 % respectively. The two control groups gave no complete responses to any of these figures.

These results conflict with the findings of Fuchs (1920), who reported the absence of completion of complex representational material. Further, they are contrary to the prediction of the Gestalt theory of perception. The completion effect could only be interpreted as the result of 'internal forces' in the perceptual organization which tend to produce simple, complete and regular configurations in accordance with the law of Prägnanz if it were restricted to simple forms. Another finding which cannot be reconciled with the Gestalt interpretation is that of completion of incomplete words across hemianopic field defects (Kinsbourne & Warrington, 1962). It was shown in six patients with left homonymous hemianopia that, when letter groups which could plausibly be regarded as representing ends of words were briefly exposed to the right of the fixation point, the tendency to the completion of geometrical figures was correlated with a tendency to respond with complete words. Purely imaginary word-beginnings seem to have been deduced from the 'ending' supplied and would appear to represent the first suitable word to come to mind.

Poppelreuter's interpretation of completion as the 'Vorstellungsmässige Ergänzung' (imaginative completion) is in part compatible with the present findings. Though he failed to account for instances of absence of completion (see Warrington, 1962) his emphasis on the importance of expectation and past experience is consistent with the present results.

A higher incidence of completion occurred among the figures which could be most readily named, although the ratio of naming to descriptive responses was much the same in the three groups of subjects. If it is accepted that readily named figures are the most familiar through experience, it may be argued that it is by virtue of their familiarity that these figures are so readily completed.

When a patient with a hemianopic field defect is briefly shown a figure adjacent to an area of defective field, he has no way of telling whether the figure in fact extended into that part of the visual field. In this respect he differs from the normal, who has available definite information of the incompleteness of figures similarly exposed. The hemianopic patient will be influenced in what he reports by two considerations: lack

of definite information as to the completeness of the figures, and expectation based on previous experience that familiar shapes appear in complete rather than incomplete form. Depending on which of these considerations is given more weight, a complete or incomplete response will result. A decision to complete is likely to be related to previous experience of the figure in its whole form rather than to its configuration and the extent to which it approximates to a 'good' Gestalt.

I should like to thank Dr Eliot Slater and the Institute of Neurology, the National Hospital, Queen Square, for facilities to conduct this study; the physicians of the National Hospital for permission to study cases under their care; and Mr A. H. Prickett for preparing and photographing the figures. I am particularly indebted to Prof. O. L. Zangwill for advice and encouragement at all stages in this work.

REFERENCES

FUCHS, W. (1920). Untersuchung über das Sehen der Hemiamblyopiker. I. Verlagerungserscheinungen. *Z. Psychol.* **84**, 67–169.

HELMHOLTZ, H. VON (1924–25). *Helmholtz's Treatise on Physiological Optics*, ed. J. P. C. Southall. Optical Society of America.

KINSBOURNE, M. & WARRINGTON, E. K. (1962). A variety of reading disability associated with right hemisphere lesions. *J. Neurol.* **23**, 339–44.

KOFFKA, K. (1935). *Principles of Gestalt Psychology*. London: Kegan Paul.

POPPELREUTER, W. (1917). *Die psychischen Schädigungen durch Kopfschuss im Kriege 1914–16* (die Störungen der niederen und höheren Sehleistungen durch Verötsungen des Okzipitalhirns). Leipzig: Voss.

WARRINGTON, E. K. (1960). The completion of visual forms across hemianopic field defects. Unpublished Ph.D. Thesis, London University.

WARRINGTON, E. K. (1962). The completion of visual forms across hemianopic field defects. *J. Neurol.* **25**, 208–17.

(*Manuscript received* 16 *December* 1964)

Brit. J. Psychol. (1965), 56, 4, *pp.* 455–460
Printed in Great Britain

VISUAL IMPERCEPTION IN PSYCHOTIC CHILDREN

By B. HERMELIN and N. O'CONNOR

Medical Research Council, Social Psychiatry Research Unit, Institute of Psychiatry, Maudsley Hospital, London, S.E.5

Psychotic speaking and non-speaking children and subnormal controls were compared on visual discrimination tasks. Tasks which could be solved through the use of kinaesthetic, brightness or size cues did not differentiate between the groups, while those which involved shape differences and directional orientation did.

One of the most characteristic features in the behaviour of many children diagnosed as 'psychotic' or 'autistic' is their apparent imperception of visual or auditory stimuli in the absence of any observable sensory defect. However, few controlled observations are available concerning which aspects of stimulus displays evoke responses in such children.

For subnormals, Zeaman & House (1963) have shown that discrimination between two-dimensional displays is more difficult than when three-dimensional stimuli are used. They also found that objects differing in shape, size and colour were easier to discriminate than stimuli differing in shape alone. Gillies (1965) has confirmed these latter findings with psychotic children.

The first experiment to be described compared subnormal and psychotic children with and without verbal comprehension on three discrimination tasks. The first of these was concerned with discrimination of position. In the second the stimuli were a patterned and an unpatterned surface, and in the third different directional orientation of otherwise identical figures had to be discriminated.

As the groups differed in their verbal ability, the question arises of whether or not linguistic impairment is directly associated with a more general imperception, which may include visual as well as auditory stimulus dimensions. It could be argued that, while some aspects of visual discrimination are directly associated with language or other mediating processes, others are not. Thus Hebb (1958) stated that there is on the whole little difference between lower and higher mammals in size and brightness generalization. On the other hand a chimpanzee could not transfer a learned response to a triangle to a triangular arrangement of circles, though a two-year-old human child could do so. He argued that this implies a difference in mechanism, as pattern perception is more dependent on experience and learning than perception of brightness or size. These mechanisms may be differentially impaired in psychotic children, and the degree of their respective association with linguistic functions may also differ. This hypothesis was tested in the second experiment to be reported.

Experiment I

Method

Subjects. Thirty children took part in the first experiment. Twenty of them had been diagnosed as psychotic, and their behaviour had been checked against a list of nine diagnostic behaviour items for childhood psychosis, which has been developed by Creak (1961). No child was

included in the experimental group unless he showed at least four of the nine behaviour manifestations in addition to the clinical diagnosis. The group included boys and girls aged between 7 years and 14 years 6 months. Half the psychotic children were residents in a hospital for childhood psychotics, and the rest, as well as the ten subnormal controls, came from four different mental deficiency institutions. None of the controls had positive scores on any of the items of Creak (1961). Selection of subjects was based on behavioural criteria only, as neurological and biochemical information proved insufficient and unreliable.

The children were matched individually on their 'mental age' scores on the Seguin Form Board. This test item, which required the placing of different shapes into their appropriate cut-outs in a board, provided an indication of perceptual–motor ability. Mental ages, when mentioned, should be regarded as such an indication rather than as an estimate of general intelligence.

The subjects were also tested on the Peabody vocabulary test. This was done by requiring the child to identify an object named by the experimenter from a display of four adjacent pictures. The psychotic children were subdivided into two groups of ten, one containing those who had obtained a minimum comprehension score of M.A. 2 years 6 months, and another including children who did not obtain such a score. This subdivision reflected the speaking ability of the children. While those with a minimum comprehension age of 2 years 6 months possessed a limited amount of useful speech, only two in the other group spoke a few words, whereas the other eight did not speak at all. All three groups were matched for chronological age, and the mean chronological and mental ages are given in Table 1.

Table 1. *Mean chronological and mental ages (in years and months) for speaking and non-speaking psychotic children*

Group	n	Chronological age		Seguin mental age		Peabody mental age	
		Mean	Range	Mean	Range	Mean	Range
Speakers	10	11–2	9–7 to 14–6	5–1	4–7 to 8–2	3–10	2–10 to 6–2
Non-speakers	10	10–9	7–2 to 14–6	6–1	3–1 to 12–0	1–11	1–8 to 2–4
Subnormal controls	10	10–5	8–5 to 13–6	5–10	4–0 to 13–6	4–2	2–4 to 6–1

Tasks. Each subject had to solve three discrimination problems. In one task (A) kinaesthetic as well as visuo-spatial cues were provided. In another task (B) the task could be solved by simply distinguishing between the total amount of brightness reflected from the surfaces of the two stimulus objects. In the third task (C) the differing orientation of two otherwise identical figures had to be perceived.

Procedure. For task A the subject was presented with two identical cardboard boxes, three inches by three inches, which had a shiny white surface. One of them was placed upside down on a table at which the child was seated, either at his right or his left side. The other box, on the opposite side, was placed on a small stand, 12 in. high, which stood on the table. A sweet was found by lifting the 'correct' box, which was either the one on the higher or lower plane for alternate subjects. Thus by reaching either 'up' or 'down' kinaesthetic as well as visual cues were provided.

For task B two boxes of the same size as in task A were both presented on the same level, one at the left and the other at the right. One box had a plain, shiny white surface. A black arrow 2 in. long and ½ in. wide, pointing up or down for alternate subjects, was painted on the second box. The 'correct' box was the one with the arrow on it and its choice was rewarded. Thus in this task the discrimination had to be made between a white ground with or without a black figure on it.

In task C both boxes were again at the same level. One was marked with an arrow pointing vertically downwards and the other with an identical arrow pointing upwards. The consistent choice of either the upward- or downward-pointing arrow was rewarded. Thus in this task the spatial orientation of two otherwise identical shapes was the discriminating feature. The order of presenting tasks A, B and C, and the right or left positions of the correct box from trial to trial were randomized according to a balanced design. For alternate subjects either the 'up' or 'down' position was correct for all three tasks. A maximum of sixty trials in three sessions

of twenty was given for each task, so that each subject received a maximum of 180 trials. A criterion of nine correct responses out of ten was adopted for regarding the tasks as learned. As we were concerned with the perception of the correct cue rather than with its discovery, the correct solution was repeatedly demonstrated to the subjects at an early point in the trial series.

Results

The number of trials before a criterion run of ten responses, of which nine were correct, was taken as the score in each case. Scores are given in Table 2. As correlated and uncorrelated data had to be compared a 'mixed design' analysis of variance was used (Maxwell, 1958). In this the appropriate error term for main effects between the groups is the 'between people within groups' term, whereas the residual is used to test the remaining effects. An analysis of variance compared psychotic speaking, psychotic non-speaking and subnormal children (Table 3). A highly significant interaction between groups and treatments was found. Subsequent t tests within groups showed that for the controls there was no difference in the number of trials needed to solve any of the three problems. For the speaking psychotic children, task C was significantly different from A, but task B differed significantly neither from A nor from C. For the non-speaking psychotic children on the other hand, every task was significantly different from each of the others. Task A was easiest, C hardest.

Table 2. *Trials to criterion (mean and* S.D. *values) for the three discrimination problems of Expt. I*

Groups	n	Conditions: No arrow (A)	One arrow (B)	Two arrows (C)
Subnormal	10	0·5 ± 0·975	5·3 ± 6·5	12·4 ± 19·5
Psychotic speakers	10	7·9 ± 9·1	18·4 ± 19·5	25·5 ± 19·5
Psychotic non-speakers	10	7·1 ± 8·45	36·1 ± 19·5	60·0 ± 19·5

Table 3. *Analysis of variance for Expt. I*

Source	S.S.	D.F.	M.S.	F	P
Between groups	12,217·69	2	6108·84	18·41	<0·001
Between treatments	11,337·63	2	5668·81	26·37	<0·001
Groups × treatments	4,982·7	4	1245·67	5·79	<0·001
Between people within groups	8,958·9	27	331·81		
Residual	11,609·70	54	214·99		
Total		89			

As far as the differences between the groups are concerned, all groups performed equally well on task A. There was no difference in the number of trials needed for any of the three tasks between the speaking psychotics and the controls. However, psychotic non-speakers differed significantly from psychotic speakers, as well as from controls on tasks B and C. This last point can be illustrated by looking at the number of subjects who solved each problem. Everybody solved problem A. In task B all the subnormal children and nine out of ten speaking psychotics were successful, while four out of ten non-speaking psychotics failed to learn. In task C one control, three speaking psychotics and all ten non-speaking psychotics were unsuccessful over sixty trials.

Thus the number of trials needed to learn each of the three tasks does not differ significantly for the subnormals, while for the speaking psychotics it is more difficult to respond to one of two differently oriented arrows. However, the most striking result is that non-speaking psychotic children differ more from speaking psychotic children than the latter differ from subnormal controls. None of the non-speaking children learned to discriminate between the two arrows. As can be seen from the results on the other two tests, this was clearly not due to lack of motivation or co-operation. It must also be remembered that these children performed as well on the Seguin Formboard test as did the other subjects. The inability to distinguish the two arrows may be related to the 'rotation phenomenon' frequently reported for brain-damaged subjects (Shapiro, 1953). Alternatively, it may be due to an inability to focus attention on the relevant cues (i.e. the point of the arrows), or to a more general imperception. Whatever the relevant mechanisms, the findings point to a relation, though not necessarily a causal one, between verbal disability and the inability to distinguish differing directional orientation of otherwise identical figures.

EXPERIMENT II

In order to investigate the relevance of some of these factors for the results, a second experiment was carried out. This was concerned with the question of whether the directional orientation, differentiating the two arrows, was particularly difficult for psychotic children to perceive, or whether there was a more general imperception of many visual dimensions.

Method

Tasks. Visual discrimination of four different dimensions was investigated: albedo, size, shape and direction. For each of these dimensions two examples were used: one which presented the variables in a concrete and the other in an abstract form. The material consisted of black line-drawings on white 3 in. by 5 in. cards. Albedo was illustrated by a black and a white square or a black and a white apple. A large or small circle and large or small ball were used in size discrimination. For shape, a triangle with straight or curved sides, and houses with roofs of the same triangular shapes were used, while a vertical and horizontal line and a standing or lying man illustrated direction.

Procedure. All children did all eight tasks, and each task was presented twenty times, giving a total of 160 trials for each child. Abstract or concrete material was presented first or second to alternate subjects and the order of presenting the four dimensions was varied, according to a 4 × 4 Latin square. Which one of a pair of stimuli presented was the correct, rewarded one was also varied alternately between subjects. The stimulus cards were placed over boxes, and the pointing to, or lifting of, a correct one was rewarded by a sweet. The placing of the rewarded stimulus card on either the left or the right side was varied over twenty trials according to a Gellerman series (Gellerman, 1933) so that the correct stimulus was presented for an equal number of trials on either side. The tasks were carried out in four sessions, with forty trials per session. The twenty autistic children who took part in the first experiment as well as four additional psychotic subjects (two speakers and two non-speakers) were used.

The children were divided into two groups of twelve according to scores obtained by them on the Peabody Verbal Comprehension test. The lowest mental age recorded on the test is 1 year 9 months, and in the first group eight subjects did not reach this limit; they were therefore assigned a mental age for verbal comprehension of 1 year 8 months. None of the other four scored higher than 2 years 8 months. The lowest scorer in the second group obtained a score of 2 years 10 months, and the highest score was 6 years 2 months. There was thus no overlap between the groups in mental-age scores on this verbal comprehension test.

The two groups were matched for chronological age (mean 10 years 11 months, range 7–2 to 14–9) and for mental age scores on the Seguin Formboard (mean 5 years 6 months, range 3–1 to 12–0).

Results

The scores were the correct number of responses in any series of twenty trials (cf. Table 4). Analysis of variance showed no difference between abstract and concrete material, but groups as well as dimensions differed significantly; none of the interactions was significant (cf. Table 5). Subsequent t tests between the dimensions showed that size and albedo did not differ in the ease with which they were discriminated, while discrimination of shape was significantly more difficult and discrimination of direction significantly more difficult still. The difference between the groups shows that children who scored higher in the verbal comprehension test did better than those with very low scores.

Table 4. *Mean number of correct responses out of 20 made by the two groups of subjects in Expt. II for the eight combinations of dimensions and conditions*

	Albedo		Size		Shape		Direction	
Groups	Abstract	Concrete	Abstract	Concrete	Abstract	Concrete	Abstract	Concrete
Psychotic speakers	16·7	16·6	17·0	17·8	15·4	16·1	13·3	14·8
Psychotic non-speakers	12·7	11·5	10·6	11·4	11·1	10·7	9·8	11·1

Table 5. *Analysis of variance for Expt. II*

Source	S.S.	D.F.	M.S.	F	P
Between groups	1,135·88	1	1,135·88	20·266	$<0{\cdot}001$
Between conditions	9·63	1	9·63	1·370	N.S.
Between dimensions	131·68	3	43·89	6·033	$<0{\cdot}001$
Groups × conditions	3·80	1	3·80	$<1{\cdot}000$	N.S.
Groups × dimensions	47·77	3	15·92	2·232	N.S.
Conditions × dimensions	27·19	3	9·06	1·277	N.S.
Groups × conditions × dimensions	3·33	3	1·11	$<1{\cdot}000$	N.S.
Between people within groups	1,233·12	22	56·05		
Residual	1,110·97	154	7·21		
Total		191			

Discussion

In both experiments reported in this paper, speaking and non-speaking psychotic children differed in their visual discrimination ability. Our observations seem to suggest that this difference is not simply due to the fact that the relevant cues could be named by the speaking group. Though some of these children were able to report verbally why they had selected one particular stimulus card consistently, most were quite unable to do so. Nevertheless they were often able to discriminate visually while many of the non-speakers did not seem to 'see' the figures on the stimulus card, though they did 'look' at them. This may be relevant to the differences between the groups on the verbal comprehension test. Two alternative though not mutually exclusive interpretations for these test results could be suggested. First, the non-speaking children may not have understood the verbal stimulus cues given to them. Alternatively, or additionally, they may have been unable to discriminate between the four adjacent pictures from which they had to select the correct response.

The gradient of correct responses from albedo and size to form and direction confirms the findings of the first experiment that directional cues are particularly

difficult for psychotic children to follow. Teuber (1959) has shown that in adults spatial integration can be disorganized by lesions in the frontal and parietal, as well as in the occipital region of the brain. Though he has also shown that brain damage in adults or children may affect performance differently, the scores on the Seguin Formboard test and the ability to select one of two identical boxes in different positions as well as the capacity to perform certain form discriminations by touch alone (Hermelin & O'Connor, 1964) demonstrate a relatively intact kinaesthetic feedback system in these children. It remains to investigate whether interference with the information coming from this system to the visual mechanism will affect the subjects. Alternatively the inability to utilize schemata and maps which is found in the parietal syndrome may be associated with the low scores obtained by the children on the vocabulary test. Verbal as well as visual discrimination ability may be relevant in this context, and may be relevant in functions additional to those which have been tested in these experiments. It could be suggested that similar underlying processes are involved in aspects of both visual and verbal discrimination. These might include cross-modal coding, short-term memory, figure–ground distinction and signal–noise ratios. Impairment in some or all of such functions may result in visual as well as verbal imperception. Brightness and size discrimination might be less affected by such imperception than discrimination of shape or direction. According to Hebb (1958) the former two do not depend on learning and experience, but the two latter do. While the mechanism for size and brightness discrimination is probably sensory, mediating processes seem to intervene in discrimination of shape and direction. Verbal ability would therefore be expected to be associated to a greater extent with discrimination than with brightness and size. This is in accordance with the present results.

We are grateful to Dr G. O'Gorman, Dr J. Gibson and Dr A. Shapiro and the staffs of Smith Hospital, St Lawrence's Hospital and Harperbury Hospital for making this study possible.

References

CREAK, M. (1961). Schizophrenic syndrome in children. *Lancet*, ii, 818.

GELLERMAN, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.* **42**, 207–8.

GILLIES, S. (1965). Discrimination learning in psychotic children and subnormal controls. *Bull. Brit. psychol. Soc.* **18**, no. 59, 7A.

HEBB, D. O. (1958). *A Textbook of Psychology*. London: Saunders.

HERMELIN, B. & O'CONNOR, N. (1964). Cross modal transfer in normal, subnormal and autistic children. *Neuropsychologica*, **2**, 3.

MAXWELL, A. E. (1958). *Experimental Design in Psychology and the Medical Sciences*. London: Methuen.

SHAPIRO, M. B. (1953). Experimental studies of a perceptual anomaly. III. The testing of an explanatory theory. *J. ment. Sci.* **99**, 394–409.

TEUBER, H. L. (1959). *Some Alterations in Behaviour After Cerebral Lesions in Man. Evolution of Nervous Control*. Washington, D.C.: American Association for the Advancement of Science.

ZEAMAN, D. & HOUSE, B. J. (1963). The role of attention in retardate discrimination learning. In *Handbook of Mental Deficiency*, ed. N. R. Ellis. New York: McGraw-Hill.

(*Manuscript received 17 March 1965*)

Brit. J. Psychol. (1965), 56, 4, *pp.* 461–465
Printed in Great Britain

PREFERRED HANDEDNESS AS A FACTOR IN THE ABILITY TO FORM A SET BASED ON IMAGINAL SITUATIONS

BY STEWART SCHWARTZ AND GILBERT SPITZER
City College, City University of New York

An experimental study is summarized in which the ability to form a set based on imaginal situations was studied under two conditions: when the imaginal situation is in direct conflict with the preferred handedness of the individual, and when the imaginal situation conforms to the preferred handedness of the individual. A fixated set resulting from the imagined situation of handling objects differing in size manifested itself in the subjects' subsequent perception of objectively equal objects or balls as being unequal. Set was more readily established in those subjects whose handedness conformed with the imaginal situation.

One of the most important determinants of an individual's response to a situation is a factor which has been variously termed mental set, preparatory set, attitude, readiness, and determining tendency. Set tends to facilitate the subject for responses he is prepared for, and tends to inhibit any competing responses. Learning is guided by a total attitude or set of the organism. The positive value of a set is due to its facilitation of appropriate responses and its inhibition of inappropriate ones. Its disadvantages appear when it does the reverse because it is not orientated to the given situation. Man has the ability to act not only in conformity with an actual situation but with an imaginary one as well, for example, a preconceived situation that has not yet materialized.

This assumption was tested in an experiment by Natadze (1960), who set out to ascertain the possibility of inducing a set on the basis of some pertinent situations which were conceived, not perceived. In six different experimental situations the subject was asked to imagine two objects (circles, weights, cylinders or balls, depending on the experiment), one on either side of him, and to imagine the object on his right as being heavier or larger than that on his left. The observer was then confronted with two equal objects, and asked which was larger, or heavier. Natadze said that when the observer responded by saying the object on the right was heavier the observer showed an assimilative set. When the observer responded by saying the object on the left was heavier or larger, he demonstrated a contrast set. Natadze in his experiment called a response in either direction a response that conformed to a set. One of the difficulties with Natadze's experiment is that he proved his hypothesis correct by his definition of set. Natadze's conclusion was that the problem of the feasibility of evolving a set on the basis of imagination had been solved.

The present experiment re-tests the ability to evolve the formation and fixation of a set on the basis of imagination using experimental situations similar to that of Natadze. If the observer's response did not conform to the imaginal situation, as stated by the experimenter, we have said that a set was not established, thereby correcting for the ambiguity in Natadze's experiment. When the response of the observers conformed to the imaginal situation we have said that a set was established.

This experiment deals more specifically with the effect of the subject's preferred-handedness in determining the set in the imaginal situation. (cf. also Rees & Israel, 1935).

Purpose. The aim of the experiment was to study the ability to form a set based on imaginal situations, under two conditions: when the imaginal situation is in direct conflict with the preferred handedness of the individual, and when the imaginal situation conforms to the preferred handedness of the individual. The hypothesis to be tested is that the feasibility of evolving a set on the basis of imagination is greater when the imaginal situation conforms to the preferred handedness of the individual.

METHOD

Subjects

Fifty undergraduate students, all right-handed and naïve experimental subjects.

Procedure

The subjects were divided into an experimental group and a control group, each of 25 subjects. Subjects were tested individually. Before each experimental situation they were asked to pick up a wooden block. In all cases the hand that they picked it up with conformed with their preferred handedness. (The subjects were asked their preferred handedness only after they had lifted the block. Those that were left-handed were rejected.)

Task I: materials—two rubber balls of equal size.

In the control group the subject was told to close his eyes, and place his hands on the table, palms up. He was told to picture as vividly as possible, as if he were actually experiencing the situation, that he was grasping two rubber balls, one in each hand, a large ball in the right hand and a small ball in the other. (This was repeated five times.) After the fifth time the two balls were placed in the subject's extended hands for a moment. He was asked to tell which ball was larger. In the experimental group the instructions were the same, except then the subject was asked to imagine the larger ball in the left hand rather than right hand.

Task II: materials—two weights of equal size and weight.

In the control group the subject was told to fix his eyes on the two weights (placed in front of him by the experimenter). The subject was then told to picture to himself as vividly as possible that the weight on the right side was very heavy (filled with lead), and that the other weight was very light (hollow). The subject was told to imagine lifting them simultaneously with both hands, the heavy one in the right hand, and the light one in the left hand. He was told to try to feel with all possible distinctness the difference in the weights. (This was repeated five times.) The subjects were then asked to lift both weights simultaneously and tell which was the heavier. In the experimental group the subjects were given the same instructions, but were asked to imagine lifting the heavier weight with the left hand.

Task III: materials—two cylinders: of equal weight, but not of equal size. (The larger cylinder was put to the right of the subject in both groups.)

In the control group the subject was told to keep his eyes on the cylinders, and picture to himself that the smaller cylinder was much heavier than the larger cylinder. He was told to visualize as distinctly as possible that he was lifting the cylinders by the attached strings, the small but very heavy one with the right hand, and the large but light one with the other. This was repeated five times. The subject was then told to hold the attached strings by the thumb and forefinger, and lift both cylinders simultaneously. He was then asked which was the heavier. In the experimental group the instructions were the same except that the subjects were told to imagine lifting the small but very heavy cylinder with the left hand.

Task IV: materials—two weights of equal size; also two cylinders of equal size and equal weight.

In the control group the subjects were told to lift two weights; after being sure that the weights are equal, the subject simultaneously was told to imagine lifting both weights and picture to himself, with the utmost vividness, that the weight in the right hand was much heavier than the one in the left, as if the one in the right were filled with lead and the one in the left were hollow. This was repeated five times. Then the subject was told to lift simultaneously the two cylinders in front of him by means of the attached strings, holding the knots with the thumb and fore-

finger, and reporting which was heavier. The experimental group was given the same instructions except that the subjects were asked to imagine the load in the left hand as being heavier than the one in the right.

Task V: materials—two cylinders of equal size but not of equal weight. (In both groups the heavier cylinder was to the right of the subject.)

In the control group the subject was told to watch the two cylinders and imagine as vividly as possible that he was lifting them by the attached strings, simultaneously, the heavy one filled with lead by the right hand and the light one (hollow) by the left hand. This was repeated five times. The subject was then asked to lift both cylinders, and report which was heavier. In the experimental group the instructions were the same except that the subject was asked to imagine lifting the heavier cylinder with his left hand.

Task VI: materials—two pieces of white paper, one with two circles of the same diameter drawn on it, the other blank.

In the control group the blank paper was placed before the subject, and he was told to imagine as distinctly as possible that he saw two circles; a large one on the right and a small one on the left. This was repeated five times. The subject was then shown the white paper with the two equal circles drawn on it, and asked which was the larger. In the experimental group the same procedure was used except that the subjects were told to imagine the larger circle on the left side of the paper.

Results

The results of this experiment are summarized in Table 1. The data show that in the control group, where the imaginal situation conformed to the preferred handedness of the subjects, the subjects more readily formed a set. In the experimental group where the imaginal situation was in direct conflict with the preferred handedness of the individual the ability to form a set was much more difficult. As the results show, it was extremely easy for the control group to form a set in task V (the two cylinders of equal size, but not of equal weight). The difference in this case is probably greater, and more distinguishable by the subjects, than in any of the other tasks. In task V the 'real' situation conformed with the imaginal one. The fact that the subjects in the experimental group for task VI had such a high percentage of 'set formed' may be explained by the fact that this experiment did not call for the use of either of the subject's hands, either in the imaginal or in the real situation.

Table 1. *Percentage of subjects showing set (and no-set)*

	Control group		Experimental group	
Task	Set	No-set	Set	No-set
I	**72**	28*	**20**	80*
II	**58**	42	**24**	76*
III	**68**	32	**20**	80*
IV	**72**	28*	**16**	84*
V	**100**	00*	**00**	100*
VI	**84**	16*	**80**	20*

* Differences significant at $P < 0{\cdot}05$ at least. For $n = 25$ there are the following correspondences between percentage dichotomies and two-tail significance levels for the binomial test against $H_0[p = q = 0{\cdot}5]$: 32 % and 68 %, $P = 0{\cdot}108$; 28 % and 72 %, $P = 0{\cdot}043$. In the 2×2 contingency table of *frequencies* for each of tasks I–V, χ^2, D.F. 1, is significant ($P < 0{\cdot}001$); for task VI χ^2, is not significant.

DISCUSSION

As 'set' is here defined; when the subjects responses conformed to the experimenter's instructions the subject demonstrated a set, when his responses did not conform to the experimenter's instructions he did not demonstrate a set. In the control group the subject's set response was that the weight (circle or cylinder) on the right was heavier or larger than the one on the left. In the experimental group the subject's set response was that the weight (circle or cylinder) on the left was heavier or larger than the one on the right. In both groups the response that the objects were equal was indicative of 'no-set' being formed. Natadze (1962) minimized the chance of the subject's response being that of equal: he did so by asking them which object was larger or smaller, which leads the subject to think that he is wrong if he gives an equal response as a choice.

Although sets were established by the method of verbal instruction, they were shown to influence the responses definitely and consistently. An interesting question would be how a group of right-handed subjects (control group) given no instructions would have responded.

Natadze induced set (in most cases) after stating the imaginal situation to the subjects fifteen times. In our experiment we stated the imaginal situation only five times. Fifteen times might have been enough to have had a hypno-suggestive effect on the subjects. Even with five times, many subjects felt as if they were being 'indoctrinated'.

Often set is visible as a posture of readiness, like that of the runner poised for a quick start. In the present experiment many subjects showed physical movement, e.g. when told to imagine holding a heavy ball in one hand, many subjects lowered that hand, or in some way moved it; some subjects were conscious of this movement, others were not. However, it seems that even in cases where the 'posture' or movement might not be observable externally, there occurred some type of internal 'posture' of muscular tensions, since, even of the subjects in whom no physical movement was observable, many complained of muscular fatigue. The intensity of the fatigue or tiredness may depend on the person's ability to perform the task, in this case the ability to imagine a situation.

So we see that it may not only be muscular fatigue that was demonstrated by many subjects, but also mental exertion which resulted from the difficulty in producing the vividness of an image. The idea that an image may evoke a set more readily if the subject has a specific relation to it seems to be acceptable on the basis of the present experiment. Imaginal situations where the subject is lifting something with his preferred hand may have more and greater vividness for him.

The time-interval after the subject is asked for his response seems to be of critical importance. Many of the subjects visualized a transposition of the illusion as the time-interval increased. Thus the set response which was given by the subject soon after he was asked to respond seems more stable than when the subject hesitated. However, a fixated set developed on the basis of imagination seems as a rule to be less stable than are sets developed on the basis of perception.

According to the data obtained in this experiment the fixated set evolved on the basis of an imagined situation is characterized by a number of peculiarities which

distinguish it from the manifestations of a fixated set evolved on the basis of perception. The data also show that the feasibility of forming a set based on an imaginal situation is greater when the person has a specific relation to the thing visualized, in this case the preferred handedness of the subjects.

The authors wish to thank Prof. Gertrude Schmeidler, under whose supervision this experiment was conducted at the City College of the City University of New York, May 1963.

References

Natadze, R. (1960). The emergence of set based on imaginal situations. *Brit. J. Psychol.* **51**, 237–45.

Natadze, R. (1962). On the question of the role of set in the perception of size constancy. *Vopr. Psikhol.* **8**, no. 5, 45–56.

Rees, H. J. & Israel, H. C. (1935). An investigation of the establishment and operation of mental set. *Psychol. Monogr.* **46**, 1–27.

(*Manuscript received* 7 *April* 1965)

Brit. J. Psychol. (1965), 56, 4, pp. 467–470
Printed in Great Britain

A SIMPLE LABORATORY DEMONSTRATION OF SUBCEPTION

By GERALD M. MURCH
Universität Göttingen, West Germany

A simple, replicable demonstration of subception is described, using a discriminatory response as the independent variable. Stimuli were exposed to thirty-six subjects in a tachistoscope until they were able to report the perception of a black line, appearing in the centre of a white field. The subjects were then asked to choose one of two geometrical figures from six pairings of four figures, each pair shown separately. The tachistoscope projections, presented in a counterbalanced design, consisted of one of the figures to be discriminated drawn in red, divided by a 3 cm black line. The black line, having a stronger contrast, was perceived before the red figure, defining the supraliminal and subliminal thresholds as the contrast relation between red and black lines. A significant tendency was found for selection of the projected figures, without verbal report of the subliminal stimulus. In a second part of the experiment one figure was to be chosen, after each projection, from the finite population of four figures; the results did not differ from chance expectation.

The use of a perceptual discriminatory response for the recovery of a subliminal stimulus was discussed, and the plausibility of a threshold for intensity relations was demonstrated.

Studies reporting or disputing the discrimination of subliminal stimuli have become increasingly numerous in the psychological literature. Many writers have either accepted or rejected subliminal stimulation as a perceptual phenomenon from the results of a relatively small number of experiments. In spite of this work there remains some confusion of terminology which could be clarified by classifying the published experiments into the two major methods of stimulus presentation that have been used: (1) subliminal perception, the presentation of both a supra-liminal and a subliminal stimulus (i.e. the subject is unaware of the subliminal stimulus); (2) subception, the presentation of only a subliminal stimulus (see Murch, 1964).

This study is concerned with the design of an experiment in subception. Early experiments in this area usually dealt with verbal responses to subliminal visual stimuli (e.g. Sidis, 1898; Poetzl, 1917; Williams, 1938). Lazarus & McCleary (1951) reported a subception effect with GSR as the independent variable, after which many studies of this type appeared (e.g. Lowenfeld, Rubenfeld & Guthrie, 1956*a*, *b*; Eriksen, 1956*a*; Wall & Guthrie, 1959). This form of response measurement has been criticized by Eriksen (1956*b*), who said that results were determined by a partial correlation between subliminal stimuli and GSR with the verbal response held constant. Dixon (1958) criticized this form of design also; 'the GSR and recognition thresholds cannot be compared due to the number of different response categories by which they are measured' (p. 212). In addition Howes (1954) presented some statistical considerations of these GSR response measurements, offering an alternative statistical explanation of the results.

The use of verbal response categories has been criticized by the present author (Murch, 1963). If the definition of the perceptual threshold is the verbalization of a perceived stimulus, a further verbalization of a stimulus presented at a lower threshold cannot be expected. It was found (Murch, 1965) that a perceptual discriminatory response between two alternatives is superior to a verbal response, the two

alternatives including the previously subliminally presented stimulus and an equally likely alternative.

A. E. Edwards (1960) criticized the 'subception effect' when the threshold is established by one method and the stimulus recovered by another, more sensitive method. The important variable is, in the opinion of the present author, the relation between the subject's ability to make a verbal report of the stimulus and his level of awareness. The subject reports the amount of information available to him. If the subject is not capable of telling the experimenter that he has seen something (verbalizing the subliminal stimulus) but decides between two choices (discrimination in a forced choice situation) above a chance level, it is then possible to speak of a 'subception effect'.

Finally, the complicated designs of many of the previous experiments can be criticized. For all practical purposes, these designs are not replicable. The reader is forced to accept the data, hoping that the designs were objectively presented and free from 'intervening variables'. The present study is an attempt to present a simple and easily replicable demonstration of subception, with a basic model for the detection of a subliminal stimulus, using a perceptual discrimination as response variable.

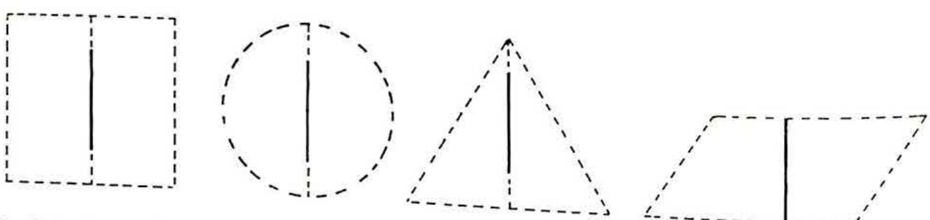


Fig. 1. The four stimulus figures. Lines which were continuous *red* lines are shown dotted; each figure had a vertical *black* line, 3 cm long.

METHOD

Four geometrical designs (circle, square, triangle and parallelogram) were prepared on 12 × 17 cm white cards to be presented in a 'Bettendorf-Electronic' two-field mirror tachistoscope. Each figure was centrally divided by a vertical line, and each was drawn in red pencil (hard). From the centre point of the vertical dividing line, a black line was drawn extending 1·5 cm above and below the entire point of each figure (see Fig. 1).

Nineteen male and seventeen female advanced semester students or members of the staff at the Institut für Psychologie in Göttingen served as subjects. Each subject was told that he would receive projections in the tachistoscope of increasing duration (1, 1·2, 1·4 msec, etc.) until he was able to identify a black line appearing in the centre of the blank second field. As soon as it had been identified the subject was to choose between two geometrical figures, drawn in red pencil, which were presented on white cards beside the tachistoscope The black line was the same in length, width and density for all figures. The six pairs of figures included all possible pairs of the four figures presented subliminally. Each subject received six series of projections, and discriminated between all six pairs of figures.

The threshold was defined as the verbal report of the black line, without report of the red figure of lesser contrast. In cases where the red figure or parts thereof were reported, the subjects were excluded.

In the second part of the experiment the subjects were first shown the four figures, presented on two cards outside of the apparatus, and told that there would be four further projections. After each report of perception of the black line one of the four figures was to be selected so that each figure could be chosen only once, and by the fourth projection the subject had but one figure left to choose

Results

The frequency with which the projected figures were selected was computed for all thirty-six subjects, giving 148 as the observed frequency of choice of the subliminally projected figure; this figure was compared with the expected frequency (108 out of 216 projections for $P = 0{\cdot}5$) by means of the normal curve test (A.L. Edwards, 1960, p. 46). This test resulted in a z-score of 5·31 ($P < 0{\cdot}001$), leading to the rejection of the null hypothesis that the subjects were selecting the figures on a random basis.

The observed frequencies for male subjects (77 out of 114 projections) and for female subjects (71 out of 102 projections) were also compared; the difference favoured the female subjects, but was not significant ($z = 0{\cdot}33$).

Table 1. *Differences between expected and observed frequencies in the selection of the subliminally projected figures*

Number of selected projections ...	4	2	1	0	D.F.	χ	P
P	0·042	0·500	0·250	0·208	3	2·67	> 0·05
expected frequency	1·31	18·0	9·00	7·49	—	—	—
observed frequency	3·00	17·0	8·00	8·00	—	—	—

Table 2. *Differences between expected and observed frequencies in the initial figure selected*

Figure	Circle	Triangle	Square	Parallel.	χ^2	P
P	0·25	0·25	0·25	0·25	12·22	< 0·05
Expected frequency	9	9	9	9	—	—
Observed frequency	10	17	6	3	—	—

For the second part of the experiment, observed and expected frequencies were compared by a χ^2 test (A. L. Edwards, 1960, pp. 63–5), but the differences were not significant (see Table 1). It appeared that the subjects were biased on the basis of previous experience or predisposition in the selection of the first figure. To test this possibility, the observed and expected frequencies for the first selected figure were compared by a χ^2 test (cf. Table 2). The significant difference shows that a bias was present in the first selection.

Discussion

It appears that the subjects, although unable to indicate discrimination by verbal report, were able to select the projected figure beyond chance expectation. In the second part of the experiment the differences were not significant. In this design a response restriction was initially placed on the subjects, so that the first response was made between four categories, which resulted in a predisposition of choice. The significant value for the first choice, showing the triangle and circle to be chosen more frequently, indicates the presence of an uncontrolled dependent variable, influencing the subjects' perception of the available response categories. The subliminal stimulation appears, in the presence of a predisposition to the stimulus material, to have no effect, at least when the response items (figures) are presented *before* the subliminal stimulus. In the first part of this experiment, the categories (figures) were first presented *after* the presentation of the subliminal stimulus.

A special method of threshold determination was used here, which appears feasible and useful in subliminal experimentation. The relation between contrasts (intensity) for the red and black lines on a white background defines the below-threshold and verbal threshold as the verbal report of the dark line, which, for all subjects, can be assumed to remain in a constant relation to the below-threshold red line.

Sophisticated subjects were used deliberately in this experiment, as it was believed that these subjects would be more capable of responding without preconceptions about the response categories, and, further, would be able to report any observations or information received other than the perception of the supraliminal stimulus (i.e. the black line).

A further note is needed on the subjects' responses about the information available for the discrimination. Twenty-five subjects (69·2 %) reported having had no information, but that they had simply selected one of the two figures in each pair, with no outside criteria. Six subjects (16·0 %) reported a 'closure' effect, possibly to be attributed to an after-image of the stimulus. The possibility that discriminations by other subjects were made on a similar basis, which could not be verbalized, must be considered, and further experiments are needed. Three subjects (11·1 %) reported a stroboscopic effect (movement of the black line) which could not be related to the stimulus figures. Two subjects (5·8 %) reported that the dark line varied in length, but all lines were 3 cm long.

REFERENCES

DIXON, N. F. (1958). Apparent changes in the visual threshold as a function of subliminal stimulation. *Quart. J. exp. Psychol.* **10**, 211–19.

EDWARDS, A. E. (1960). Subliminal tachistoscopic perception as a function of threshold method. *J. Psychol.* **50**, 139–44.

EDWARDS, A. L. (1960). *Experimental Design in Psychological Research* (revised ed.). New York: Holt, Rinehart and Winston.

ERIKSEN, C. W. (1956*a*). An experimental analysis of subception. *Amer. J. Psychol.* **69**, 625–34.

ERIKSEN, C. W. (1956*b*). Subception: fact or artifact? *Psychol. Rev.* **63**, 74–80.

HOWES, D. (1954). A statistical theory of the phenomenon of subception. *Psychol. Rev.* **61**, 98–110.

LAZARUS, R. S. & MCCLEARY, R. A. (1951). Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.* **58**, 113–22.

LOWENFELD, J., RUBENFELD, S. & GUTHRIE, G. M. (1956*a*). Verbal inhibition in subception. *J. gen. Psychol.* **54**, 171–6.

LOWENFELD, J., RUBENFELD, S. & GUTHRIE, G. M. (1956*b*). Stimulus generalization in subception. *J. gen. Psychol.* **54**, 177–82.

MURCH, G. M. (1963). Die Existenz des Phänomens der unterschwelligen Wahrnehmung und die Bedingungen zu einer Darstellung jenes Phänomens. Unpublished thesis, Universität Göttingen.

MURCH, G. M. (1964). Suggestion for clarification of terminology in experiments on subliminal stimulation. *Percept. mot. Skills*, **19**, 442.

MURCH, G. M. (1965). A set of conditions for the consistent recovery of a subliminal stimulus. *J. appl. Psychol.* **49**, 257–60.

POETZL, O. (1917). Experimentell erregte Traumbilder in ihrer Beziehungen zum indirekten Sehen. *Z. ges. neurol. Psychiat.* **37**, 278–349.

SIDIS, B. (1898). *The Psychology of Suggestion*. New York: Appleton.

WALL, H. W. & GUTHRIE, G. M. (1959). Extinction of responses to subceived stimuli. *J. gen. Psychol.* **60**, 205–10.

WILLIAMS, A. C. (1938). Perception of subliminal visual stimuli. *J. Psychol.* **6**, 187–99.

(*Manuscript received* 14 *December* 1964)

Brit. J. Psychol. (1965), 56, 4, *pp.* 471–475
Printed in Great Britain

ABILITY AND SEASON OF BIRTH

By J. E. ORME
Middlewood Hospital, Sheffield

The evidence available suggests that an association between season of birth and subsequent ability can occur for a number of reasons. In the present study, groups of subnormal and supernormal ability are used to examine one relevant factor. This hypothesis is that climatic temperature changes during fœtal development are associated with subsequent ability at all levels, and not only with subnormal subjects. The hypothesis is confirmed, although for climatic environments different from that of the British Isles the particular relationships between climate and ability might vary. The investigation of such relationships, however, might provide a mode of examining the processes whereby ability develops from its genetic origins.

Within a group of subnormal subjects, season of birth appeared to be associated with subsequent intellectual performance (Orme, 1963*a*). Birth in summer and autumn was significantly advantageous compared with birth in winter and spring. Crookes (1963), and Davies (1964), with subjects of more normal ability, found no evidence for such an association. They suggest that any association between season of birth and ability might be peculiar to subnormal subjects.

Pasamanick & Knobloch (1961) review work on American subjects showing a complex of causes resulting in a higher incidence of subnormal births in winter and spring. They note evidence that such a seasonal effect was itself related to conception in relatively hot summers. Two underlying causes of such a phenomenon could be a reduction of maternal protein intake or a direct stress effect operating through the hypothalmico–pituitary–adrenal–cortical axis, producing fœtal damage.

It seems reasonable to expect that if climate can have such effects, variations in climate over the world will produce varying seasonal effects. The writer's findings (Orme, 1963*a*) were concerned with variations within a subnormal group rather than in the general incidence of subnormality. Furthermore, it appeared that climatic temperature trends during the whole of fœtal development were important rather than conditions at birth or any particular time of pregnancy. It seems possible that the association between season of birth and subsequent ability includes a number of factors. Some of these are clearly pathological as they contribute to the general incidence of subnormality. Other factors seem to be of a more normal kind in that their effect contributes to variations of ability at all levels. Fitt (1941) reviewed and reported studies with subjects of normal ability, showing a season-of-birth effect, on intelligence-test performances. Furthermore, an individual's performance on such measures as rate and efficiency of work also fluctuated over the year in a seasonal cycle. Fitt thought the overall evidence suggested a basic endogenous factor of reduced endocrine activity (involving the pituitary and hypothalamic activity) in winter. Such a factor, in mammals, might be related to more obvious phenomena such as hibernation. But if the writer's findings were due to a process not peculiar to subnormals, the negative findings of Crookes and Davies require explanation. The present paper presents a hypothesis and study with this aim of rationalizing investigations into one aspect of the season-of-birth and ability relationship.

The mean I.Q. difference reported (Orme, 1963*a*) between summer and winter births is relatively small. Now if this small seasonal difference is operant at all levels of ability it can be represented by two normal distribution curves, one for summer and one for winter births, the mean values of each being fairly close together. It can then be argued that within two or even three standard deviations from the mean of the general population, the difference in frequency between winter and summer births is small compared with the absolute frequency. It is only at distances of more than three standard deviations from the mean that the seasonal difference in frequency becomes an appreciable proportion of the absolute frequency. For example, only with I.Q. of less than 55 (with S.D. = 15) will winter births clearly predominate. This is in keeping with the results of Orme (1963*a*) studying variations within a subnormal group. It would also partly account for the overall raised incidence of winter and spring subnormal births as shown by Pasamanick & Knobloch (1961). Conversely, only with I.Q. greater than 145 (with S.D. = 15) will summer births clearly predominate.

With more average ranges of ability quite large samples will be needed to demonstrate an association between season of birth and ability. In fact, a number of the relevant studies reported by Fitt used samples of several thousands. Both Crookes (1963) and Davies (1964) used samples which on the above hypothesis would be too small to demonstrate a seasonal effect at more or less average ranges of ability. But apart from these considerations the seasonal effect should directly reflect the main seasonal variable, climatic temperature. It is here, rather than in differences in proportion of incidence, that perhaps the main test lies.

If this argument is valid, the seasonal effect should be demonstrable at both extremes of the normal curve with a predictable relationship to mean changes in seasonal temperatures. Exceptions to such an association should, on examination, show an association with atypical temperature conditions, again of a predictable kind, during the fœtal development of such individuals.

Subjects

Subnormal subjects. All the subnormals were in-patients of hospitals for the mentally subnormal in the Sheffield area. They were initially assessed for routine clinical purposes other than the present study. Rather more than half were females. The age range was 15–65 yr, thus giving a good sample of climatic variations in Britain from year to year.

I.Q. was derived, as in previous studies, from the Coloured Progressive Matrices (Orme, 1961, 1962). A modification of the original table of I.Q. equivalents specially prepared by the author was used for subjects up to 25 years. The present 40–54 I.Q. sample (fourth S.D. range) is therefore larger than that used previously and the original sample that it contains is slightly different. In addition, a group who fall in the fifth standard deviation range (less than 40 I.Q.) are also included.

Supernormal subjects. The Mensa organization agreed to a sample of their members being used. Of 300 circularized, complete data were finally available for 209. All Mensa members have an I.Q. on the Cattell IIIB test (S.D. = 24) over 147 (Cattell, 1952). Such a test-performance is a prerequisite of entry and as it is carried out under formal conditions it is presumably reliable. Unfortunately, the test has a ceiling I.Q. of 161 with a young adult standardization sample. For subjects over 39, age allowances were calculated enabling a fourth standard-deviation range to be separated out. For the 19–39 age-range, however, calculations can only be made on a proportionate basis by assuming the ratio of the third to the fourth S.D. incidences to be roughly 16:1.

According to the hypothesis presented earlier, a seasonal frequency should occur between subnormal and supernormal groups beyond the third S.D. range. Furthermore there should be an even more significant association between seasonal climatic temperatures and seasonal frequencies throughout the year. This association should be negative for the subnormal group in that decreasing temperatures should go with increasing numbers of subjects. With supernormal groups, the association should be positive in that decreasing temperatures should go with decreasing numbers of subjects. The basic argument is that seasonal climatic changes are directly responsible for seasonal variations in ability (via some hypothesized channel of physiological reactivity). The major variable in seasonal climatic changes is one of temperature, precise records of which are available. In a previous study (Orme, 1963*a*), the temperature records for the 10 months of fœtal life of each subject was compared with the mean temperatures over the years. This study indicated that apparent exceptions to the seasonal prediction could be explained by relatively unusual temperature conditions during fœtal development.

Results

Table 1 gives the seasonal distribution of birth in the subnormal and supernormal groups together with the mean seasonal climatic temperatures. Considering first the groups beyond the third S.D. range, it is readily apparent that variations in seasonal incidence are closely associated with variations in seasonal temperature. As hypothesized, this association is negative with the two subnormal groups and positive with the supernormal group.

Table 1. *Seasonal variations in temperature and in incidence of birth*

	n	Winter (Nov.–Jan.)	Spring (Feb.–Apr.)	Summer (May–July)	Autumn (Aug.–Oct.)
Mean temp. °F ...		40·9	42·3	56·5	55·5
Subnormals					
5th S.D. range	54	15 (*28* %)	18 (*33* %)	8 (*15* %)	13 (*24* %)
4th S.D. range	126	38 (*30* %)	34 (*27* %)	26 (*21* %)	28 (*22* %)
3rd S.D. range	99	19 (*19* %)	23 (*23* %)	35 (*35* %)	22 (*22* %)
Supernormals					
3rd S.D. range	160	43 (*27* %)	35 (*22* %)	44 (*28* %)	38 (*24* %)
4th S.D. range	49	9 (*18* %)	11 (*22* %)	16 (*33* %)	13 (*27* %)

The three groups split over four seasons give a total of twelve subgroups for which a correlation can be calculated between incidence and seasonal temperature. For comparability, percentage incidences were used. With the supernormal group, the extent of the percentage deviation from 25 % (the expected percentage for practical purposes as indicated by the Registrar General's returns) was reversed for comparability with the subnormal groups. The product moment correlation coefficient between seasonal temperatures and incidence is $+0{\cdot}89$ ($n = 12$; $t = 6{\cdot}396$; $P < 0{\cdot}001$). An alternative statistic can be made from the fact that the order of incidence between the seasons almost perfectly follows the rise and fall of temperature in the predicted manner. The only exception, a minor one, is that with the fifth S.D. range of subnormals, the highest incidence is in spring with winter second. The probability that such an almost exact matching with expectancy could occur in three groups by chance is very small ($P < 0{\cdot}001$). Finally, although the supernormal group is relatively small, the distribution between subnormal and supernormal groups for winter and summer incidence gives $\chi^2 = 4{\cdot}22$ (D.F. 1, with Yates' correction; $P < 0{\cdot}05$).

The seasonal distributions are also included in Table 1 for the subnormal and supernormal third S.D. ranges. As expected from this paper's hypothesis, the winter and summer incidences in the supernormal group are equivalent. With the subnormal group, however, there is almost twice the frequency in summer than winter, a ratio significantly different from the observed frequency in the other subnormal groups ($\chi^2 = 7{\cdot}68$; D.F. 1, with Yates' correction; $P < 0{\cdot}01$). Such an exception to the seasonal hypothesis, as noted earlier, should be attributable, on examination, to atypical temperature conditions during the fœtal development of the deviant group. This can be calculated by determining how many months of fœtal development, for the individuals of this deviant group, were warmer or cooler than the mean climatic temperature for each month. As the seasonal effect studied here appears to be a direct consequence of normal yearly temperature trends, the relation between season of birth and ability can be affected by such temperature fluctuations. Such an argument, in fact, was demonstrated to have some statistical standing in the writer's previous study (Orme, 1963*a*). It is of interest that the only individual season that then showed such a reliable effect was summer.

Table 2. *Climatic temperature characteristics of the ten calendar months involving fœtal development, for summer births in the third deviation range of subnormal subjects*

(n = 35)

No. of months with above-average temperature for that month . . .	10	9	8	7	6	5	4	3	2	1	0
No. of subjects	0	4	6	4	3	11	5	0	2	0	0

For summer-born subnormals in the third S.D. range, Table 2 shows the number of fœtal months with a climatic temperature above the mean for that month. Each individual has the possibility of having between none or all ten of the relevant calendar months being above the average. The expected distribution in a typical seasonal series would be symmetrical about the mid-point of 5. Table 2 shows a distribution skewed above this value. In fact, the ten subjects making up this asymmetrical distribution account for the above-expectancy incidence of summer births in this range. Relatively favourable conditions, in other words, have probably reduced the number of individuals who, in average climatic conditions, would have been more severely subnormal. Alternatively, such conditions could have affected the death-rate.

Discussion

The present investigation had two major objectives. First, to demonstrate that one source of an association between season of birth and ability was normally operant throughout the population as a whole. Secondly, to establish that this particular seasonal effect was related to climatic temperature changes throughout fœtal development.

The present results offer confirmatory evidence for both these objectives. Further studies are required, both of subnormal and supernormal subjects, to investigate what appears to be one ante-natal determinant of subsequent ability. It is possible that the particular effect found here might not be operant in countries with different climatic conditions but there is growing evidence to suggest that some kind of relation between climate and ability frequently occurs. Clearly, climatic variations do have

important effects on performance, both of a permanent (McEwan, 1965) and a temporary kind (Cattell, 1957). Apart from their intrinsic interest, the ultimate objective of such studies must be the clarification of the processes involved. The detailing of such processes should throw light on the development of ability from its genetic origins.

In the field of educational attainment, in contrast to that of ability, summer birth may prove something of a handicap to a child (Williams, 1964). An undue proportion of summer-born children might be found in special schools. Such a finding appears to be a result of differential entry to infant school and to an 'age-group position' effect. Summer-born children spend less time than others at infant school and remain the youngest in their school group.

Ante-natal influences have also been claimed to affect the development of personality. Stott (1962) has produced evidence for such an influence in behaviour disorders. The particular factor there seems to be a pathological one involving fœtal damage of the kind detailed by Pasamanick and Knobloch. On the other hand, Barry & Barry (1961) show an association between winter and spring births and a raised incidence of schizophrenia. The present writer, however, has suggested that such a relationship is a reflexion of the effect on intelligence already discussed (Orme, 1963*b*). A subsequent study (Barry & Barry, 1964) provided more evidence in keeping with such an interpretation.

I wish to thank the Mensa organization and its members for their contribution to this study. Thanks are also due to the National Society for Handicapped Children for financial assistance.

References

Barry, H. & Barry, H. (1961). Season of birth. *Arch. gen. Psychiat.* **5**, 292–300.

Barry, H. & Barry, H. (1964). Season of birth in schizophrenics. *Arch. gen Psychiat.* **11**, 385–91.

Cattell, R. B. (1952). *Handbook for Cattell Intelligence Tests*, 3rd ed. London: Harrap.

Cattell, R. B. (1957). *Personality and Motivation, Structure and Measurement.* London: Harrap.

Crookes, T. G. (1963). A note on intelligence and date of birth. *Brit. J. med. Psychol.* **36**, 355–56.

Davies, A. D. M. (1964). Season of birth, intelligence and personality measures. *Brit. J. Psychol.* **55**, 475–6.

Fitt, A. B. (1941). *Seasonal Influence on Growth, Function and Inheritance.* London: Oxford University Press.

McEwan, P. J. M. (1965). Climate and intelligence. *Brit. J. soc. clin Psychol.* **4**, 8–13.

Orme, J. E. (1961). The Coloured Progressive Matrices as a measure of intellectual subnormality. *Brit. J. med. Psychol.* **34**, 291–2.

Orme, J. E. (1962). Intelligence and season of birth. *Brit. J. med. Psychol.* **35**, 233–4.

Orme, J. E. (1963*a*). Intelligence, season of birth and climatic temperature. *Brit. J. Psychol.* **54**, 273–6.

Orme, J. E. (1963*b*). Season of birth, pyschosis and intelligence. *Dis. nerv. Sys.* **24**, 489–90.

Pasamanick, B. & Knobloch, H. (1961). Epidemiologic studies on the complications of pregnancy and the birth process. In *Prevention of Mental Disorders in Children*, ed. G. Caplan. London: Tavistock.

Stott, D. H. (1962). Evidence for a congenital factor in maladjustment and delinquency. *Amer. J. Psychiat.* **118**, 781–94.

Williams, P. (1964). Date of birth, backwardness and educational organization. *Brit. J. educ. Psychol.* **34**, 247–55.

(*Manuscript received* 21 *January* 1965)

Brit. J. Psychol. (1965), 56, 4, *pp.* 477–478
Printed in Great Britain

'CROSS-OVER' EFFECT AND SUBJECTIVE INTENSITY

By A. Z. ARTHUR
University of Canterbury, New Zealand

Audley & Wallis (1964) and Wallis & Audley (1964) have described experiments in this *Journal* which demonstrated the existence of what they called the 'cross-over' effect in the speed of relative judgement. Thus, in brightness discrimination, subjects react faster to the brighter of two stimuli when both are relatively bright and to the darker one when both are dark. Correspondingly subjects respond more quickly to the higher of two pitch stimuli when they are high but to the lower one when they are low. Having established the generality of the cross-over phenomenon Wallis & Audley (1964) were concerned with its explanation. They put forward an explanation derived from a model of choice behaviour devised previously by Audley, but finally doubt its applicability.

It is proposed here that a more satisfactory explanation of the cross-over effect can be made in terms of the experienced, or subjective, intensity of stimuli. It is unnecessary to postulate cognitive factors as an additional variable. Audley and Wallis make a serious error when they fail to make allowance for the fact that variation in the physical intensity of stimuli does not necessarily correspond to variation in the subjective intensity of stimuli. Thus the most relevant theory, the Adaptation Level (AL) theory (Helson, 1959), indicates that there is no fixed reference point against which evaluations of stimulus attributes can be made. On the contrary, the reference point is constantly shifting in accordance with the influence of focal, background and residual (previously experienced) stimuli. The design of Audley and Wallis's studies was such as to give a number of different opportunities for shifts in the subjective intensity of stimuli. We shall confine our remarks to the study on pitch.

Some minor shifts, which they do not attempt to explain but which need an explanation, can be considered as due to the operation of contrast and assimilation phenomena. It appears that the strikingly faster response to the 2900 c/s stimulus which gives a V-shape to the 'choose higher' curve (Wallis & Audley (1964), p. 126) may be one example of these phenomena, though other explanations based on stimulus intensity easily suggest themselves. Another point at which they operate is at the change-over from H to L and L to H randomization blocks. When choosing higher stimuli the AL, that is, the geometric mean of the response stimuli, is close to 1155 c/s and when choosing lower stimuli the AL is closer to 931 c/s. Thus at the point of change-over some stimuli will be differentially contrasted and assimilated. Some such effects are apparent in figs. 5 and 6 (Wallis & Audley (1964), pp. 128–9).

The major feature of the data, i.e. that subjects respond faster to the higher of two high pitches, etc., is in accordance with the expectation that one responds faster to the more distinct, or intense stimulus, intensity being measured either in terms of lowness or of highness. It is expected further that the stimulus which is further from the AL is, in general, the more intense. Thus of two high stimuli the

higher is the more intense on the high side of the AL and of two low stimuli the lower is the more intense on the low side of the AL, because they are further away from the AL. Audley and Wallis fail to take account of the fact that the correct response in their instructions is always to the more intense stimulus on the appropriate side of the AL, and that this may fully account for the faster judgement times on that side of the stimulus continuum. However, the effects of stimulus intensity in their study are further complicated by the operation of time errors. Considering only responses to low stimuli, one would expect the subjective difference in intensity between the pair to be reduced when the lower is presented first (increasing RT) and increased when the lower is presented second (reducing RT). Responses to the lower stimulus when the pair are high should be faster when the first stimulus is the lower since under this condition the subjective difference between them increases. Fig. 7 (Wallis & Audley (1964), p. 130) is in accord with these predictions.

It is to be noted that in a study of weight-judgement Michels & Helson (1954) obtained a cross-over effect by varying the time-order of presenting the standard and the variable stimuli, and they attributed this entirely to time-order errors. Owing to the complex design of Wallis and Audley's studies, as compared with the usual psychophysical experiments, it is not clear to what extent the major features of their results are explicable in terms of time-order effects alone. The most important point to be made, however, is that only after the various effects of intensity on discrimination (and hence judgement time) are properly controlled is it possible to begin to look for any choice or cognitive factors such as they suggest. The cross-over effect in their experiments is neither an original discovery nor to be explained without reference to Michels & Helson (1954), who found it earlier and who have incorporated its explanation in a general theory with quantitative parameters.

References

Audley, R. J. & Wallis, C. P. (1964). Response instructions and the speed of relative judgements. *Brit. J. Psychol.* **55**, 59–73.

Helson, H. (1959). Adaptation Level Theory. In *Psychology: A Study of a Science*, ed. S. Koch. London: McGraw-Hill.

Michels, W. C. & Helson, H. (1954). A quantitative theory of time-order effect. *Amer. J. Psychol.* **67**, 327–34.

Wallis, C. P. & Audley, R. J. (1964). Response instructions and the speed of relative judgements. *Brit. J. Psychol.* **55**, 121–32.

(*Manuscript received 21 January* 1965)

Brit. J. Psychol. (1965), **56**, 4, *pp.* 479–482
Printed in Great Britain

'CROSS-OVER' EFFECT: A REPLY TO ARTHUR

BY R. J. AUDLEY AND C. P. WALLIS
University College, London, and University of Durham

In two papers recently published in this *Journal* (Audley & Wallis, 1964; Wallis & Audley, 1964), experiments were reported which showed that judgement time in simple discrimination tasks is affected by the manner in which the subject is instructed to express his judgement. It was shown that if two stimuli lie high on a particular sensory dimension then the subject is quicker at making the judgement when he is required to choose the higher of the two than when it is the lower which is demanded; conversely, for two stimuli of low value it is the choice of the lower which leads to the quickest response. This phenomenon was referred to as a 'cross-over' effect. A provisional explanation was offered in terms of theory of choosing, which postulates a competition between alternative responses.

Arthur (1965) contends that these results would be more satisfactorily explained by considering the way in which the subjective intensities of the various stimuli are affected by the total context of stimulation in the experiment. He then attempts to use adaptation level (AL) theory to account for the results and also asserts that the cross-over effect has already been reported and explained in AL terms by Michels & Helson (1954).

It is the purpose of this paper to refute his major criticisms, to clarify some points on which we believe Arthur has misunderstood or misinterpreted our arguments, and finally to show that AL theory, in at least one form in which he has used it, would predict exactly the reverse effect to that which we observed.

First, it should be noted that Arthur restates our findings in terms which appear to minimize the fact that our experiments were concerned with the speed with which a discriminative judgement was made. Thus he states that we find that 'subjects react faster to the brighter of two stimuli when both are relatively bright'. The whole of his argument appears to be based on the assumption that our subjects did not make discriminations, but rather reacted immediately to the 'absolute' value of one of the stimuli presented to them. In point of fact we were also concerned about this very issue, and in the first paper conducted two experiments (last paragraph of p. 65, *et seq.*) to determine whether in fact the cross-over effect could be attributed to a faster reaction to that member of a stimulus pair which is further from the background. The evidence was incompatible with this hypothesis, yet it is this which Arthur seeks to advance as an explanation for the phenomenon. He makes no mention of these studies but refers only to the experiments involving pitch discrimination. Even here it is difficult to accept his explanation. It would seem to be even more necessary, with four stimulus levels, for subjects to compare the two serially presented tones before choosing the one appropriate to their judgement. If the extreme member is more distinct, it should make the discrimination easier under any response set.

Consider first what would happen under assimilation theory if the subject is comparing two low tones. According to the theory, the trace of any stimulus assimilates

towards the AL. Thus, if the lowest tone is presented first its trace will increase and a more difficult discrimination will result. If the higher stimulus is presented first the trace of the first stimulus becomes higher still, making an easier discrimination. But this will be true regardless of the response set given to the subject; the essence of our finding was that the response set itself produced a marked disparity in judgement times when the appropriate choice was the second of the stimulus pair. In fact, if we ignore response set, then for both the extreme low and high pairs of stimuli there is a difference of only 0·003 sec between the average speeds of responding to the first and second member.

According to contrast theory, all stimuli are contrasted with the AL and therefore appear more extreme than they would in some more neutral context, but this effect lasts only while the stimulus is present. Therefore, for low stimuli, the second tone of the pair will momentarily appear lower than it is, and if it is indeed the lower tone this will make for an easier discrimination. Consequently, the prediction must be again that the judgement times will always be quicker when it is the second stimulus to which response must be made. So this theory also fails to explain the very large effects of the response set which were found when the extreme stimulus was second.

At one point in his paper Arthur puts forward the rather ingenious idea that a different AL should be considered when the subject is choosing the higher stimulus than when he is choosing the lower one. Arthur calculates that the appropriate values are 1155 c/s and 931 c/s respectively. This notion of two ALs would appear to us to be entirely unacceptable in a situation where a subject must attend to both members of a pair of stimuli and cannot know in advance which will be the higher or lower. It is, however, interesting to consider the consequences that follow from this speculation. When a subject is under a response set of 'Choose higher', a higher AL is postulated and therefore a high pair of stimuli are less far from the AL; and so less susceptible to assimilation and contrast effects which depend on distance from the AL. Therefore, when choosing the higher of a pair, a high pair of stimuli should be relatively more difficult, and a low pair of stimuli should be more difficult under a set to choose the lower. A cross-over effect is certainly predicted but is the reverse of that obtained!

Arthur also suggests that the cross-over effect was observed by Michels & Helson (1964) in a study of weight judgement. Their study, in any case, was not concerned with judgement time. However, it will be examined in fuller detail. The experiment required subjects to compare each member of a set of variable weights with one of two standards (400 and 350 g), giving their judgement in terms of a 9-point rating scale. In any one series of judgements, only one standard was used, consequently this is a version of the constant method and not a 'floating standard' design as was ours. In half the presentations the standard was first and the variable second (S_1V_2) and for the remainder the order was reversed (V_1S_2). Subjects were always required to judge the second stimulus in relation to the first. Their results are summarized in Table 1. It is clear that time-order errors were always negative, and did not change direction according to which standard was used; and this would have been the nearest equivalent to a cross-over effect in Arthur's terms. We would not ourselves expect a cross-over when various standards are used in *different* series.

When a range of standards is judged in the same series, time-order errors of 'central tendency' are often found, which we have suggested (Wallis & Audley, 1964, p. 132) are analogous to the cross-over effect. A central tendency results from a proneness to make too many judgements of 'lower' with relatively low standards, and too many judgements of 'higher' with high ones. This is true for both the S_1V_2 and V_1S_2 ordering of stimuli (Woodrow, 1933). As Woodrow points out, in addition to central tendency effects there are errors due to the length of time-interval between the stimulus pair, which can be called 'pure' time-error effects. Such errors are customarily positive for short intervals, and increasingly negative for longer ones. Michels and Helson have demonstrated that the size of this 'pure' time-error is unaffected by the ordering of the stimulus pair. That finding does not seem to be relevant to the 'cross-over' effect we described.

Table 1

Standard (g)	Order	PSE (or AL) (g)
350	S_1V_2	338
	V_1S_2	365
400	S_1V_2	381
	V_1S_2	417

Time-order errors are: for S_1V_2, E_1 = PSE − standard; for V_1S_2, E_2 = standard − PSE.

In conclusion, it must be said that we do not disagree with Arthur's contention that the subjective intensity of stimuli may be affected by contextual effects of various kinds. We have explicitly mentioned the fact that a subject's judgements about even a single stimulus are relative to the total set he experiences in the appropriate context (Audley & Wallis, 1964, para. 2, p. 70); and to this extent the notion of an AL is accepted in our formulation. Nor are we completely sure that some feature of the situation, such as the chosen stimulus being further from the AL, does not have some influence on the cross-over effect we observed. Yet our attempts to find this influence in the brightness studies were unsuccessful. Perhaps there are better experiments that could be used for this purpose, but at the moment our evidence is against the view that Arthur proposes. Above all, we do not consider that he has given due weight to the fact that our subjects were making judgements. Laws about the speed of reactions to stimuli in isolation cannot be applied directly to situations involving choice without bringing in further principles. We will not argue that we have given the true explanation of the cross-over effect. We can only assert that it exists and hope that further experimental evidence can be obtained to clarify its basis. Incidentally, Arthur is wrong to state in his opening paragraph that we doubted the applicability of the choice model to our findings; our doubts were directed towards the hypothesis that a subject had a natural mode of judgement and then translated the result of this into the form demanded by the experimenter.

References

Arthur, A. Z. (1965). 'Cross-over' effect and subjective intensity. *Brit. J. Psychol.* **56**, 477–8.

Audley, R. J. & Wallis, C. P. (1964). Response instructions and the speed of relative judgements. *Brit. J. Psychol.* **55**, 59–73.

Michels, W. C. & Helson, H. (1954). A quantitative theory of time-order effects. *Amer. J. Psychol.* **67**, 327–34.

Wallis, C. P. & Audley, R. J. (1964). Response instructions and the speed of relative judgements. *Brit. J. Psychol.* **55**, 121–32.

Woodrow, H. (1933). Weight discrimination with a varying standard. *Amer. J. Psychol.* **45**, 391–416.

(*Manuscript received* 29 *July* 1965)

Brit. J. Psychol. (1965), 56, 4, pp. 483–484
Printed in Great Britain

COMMENT ON EYSENCK'S ACCOUNT OF SOME ABERDEEN STUDIES OF INTROVERSION–EXTRAVERSION AND EYEBLINK CONDITIONING

By FRANK M. McPHERSON
Department of Psychiatry, University of Edinburgh

In two recent papers, Eysenck (1965*a*, *b*) cited experiments carried out at the University of Aberdeen as supporting his hypothesis relating introversion and conditionability. Eysenck (1965*b*) described two of these studies by Brebner (1957) and by Symon (1958); but his account of this work contained several inaccuracies and omissions.

Brebner (1957) compared the speed of conditioning of eight non-sensitized extraverts, and eight non-sensitized introverts (whose MPI E-scale scores ranged from 2–16, not 6 as reported by Eysenck). The measure was the number of acquisition trials required to reach the criterion of four successive conditioned responses (CRs). Analysis of variance showed that the introverts reached the criterion sooner, the difference being significant, $P < 0{\cdot}05$ (not $P < 0{\cdot}01$, as reported by Eysenck). However, this finding must be qualified in two ways. In the analysis, the interaction between sex of subjects and extraversion was also significant ($P < 0{\cdot}05$). The difference due to extraversion alone must therefore be interpreted with caution. Moreover, although noting this interaction, Eysenck omitted to mention that there was no significant difference between the female introverts and female extraverts. Only with the male subjects was the difference significant ($P < 0{\cdot}01$). This sex difference is not predictable from Eysenck's hypothesis since the female subjects had slightly more extreme MPI E-scale scores than the males. The mean MPI E-scale scores were, for the introvert group, 11·75 (males) and 9·25 (females), and for the extraverts 37·5 (males) and 39·75 (females).

A second point is that Brebner included in his analysis only those non-sensitized subjects who had conditioned to the criterion. Subjects who had failed to condition after 110 acquisition trials were excluded. There were three such subjects, *all of whom were introverts* (MPI E-scale scores 8, 9 and 15). Thus, although no extravert failed to condition, three out of eleven non-sensitized introverts did. Eysenck (1965*a*, *b*) has not referred to this finding, which is difficult to account for in terms of his hypothesis.

Symon (1958), in the second Aberdeen experiment, followed Franks (1956) in using, in addition to the acquisition score referred to by Eysenck, an extinction score based on the number of CRs made during the ten extinction trials. There was no significant difference between the introverts and extraverts on this latter measure. Indeed, Symon noted that 'towards the latter half of the extinction trials, the extraverts were tending to make more CRs than the introverts' (p. 47). Franks, of course, had found a highly significant difference on this measure.

On the acquisition score, Symon's analysis of variance showed that the introverts

made more CRs ($P < 0{\cdot}05$). However, this analysis was carried out on the raw scores. Inspection and testing of the data showed that these scores were not suitable for analysis since the requirement of homogeneity of variance was not satisfied. The writer transformed these scores to logarithms and repeated the analysis of variance. The difference between the transformed scores of the introverts and extraverts was not significant ($F = 3{\cdot}29$; D.F. 1, 12; N.S.).

In addition to Brebner's and Symon's experiments, the present writer had carried out a third Aberdeen study which was not referred to by Eysenck in either of his papers. McPherson (1960) used an experimental design that was similar in all respects to design used by Symon (1958). The subjects were the extreme groups of 150 students. The eight introverts had MPI E-scale scores of 4–14, and the eight extraverts had scores of 36–44. On the acquisition trials, the extraverts made more CRs than the introverts (64 and 16 respectively), as they also did on the extinction trials (33 and 11). The extraverts who conditioned also made their first CRs very slightly earlier (mean for six subjects = 4·13 trials) than did the introverts (mean for five subjects = 4·50 trials). None of the differences was statistically significant. The analysis of variance carried out on the combined Symon and McPherson data also showed no significant result.

Thus, the Aberdeen work appears to show little evidence that introverts are more conditionable. On all three measures used by McPherson (1960), the extraverts conditioned better. The difference between the groups was not significant for two of Symon's (1958) measures—acquisition and extinction scores; although introverts did make more CRs. In Brebner's experiment, (1957) the sex difference, the failure of some introverts to condition at all, and the conditioning of all the extraverts are difficult to account for in terms of the hypothesized relation between introversion and conditionability. An explanation of these and other Aberdeen findings in terms of anxiety aroused by the experimental situation was suggested by Reid (1964), under whose supervision all three Aberdeen studies were carried out.

REFERENCES

BREBNER, J. M. T. (1957). An experimental investigation of the relationship between conditioning and introversion–extraversion in normal subjects. Unpublished M.A. thesis, University of Aberdeen.

EYSENCK, H. J. (1965*a*). Personality and conditioning: the present position. *Proc. Annu. Conf. Brit. Psychol. Soc., Aberdeen*, 1965, *Bull. Brit. psychol. Soc.* **18**, no. 59.

EYSENCK, H. J. (1965*b*). Extraversion and the acquisition of eyeblink and GSR conditioned responses. *Psychol. Bull.* **63**, 258–70.

FRANKS, C. M. (1956). Conditioning and personality: a study of normal and neurotic subjects. *J. abnorm. soc. Psychol.* **52**, 143–50.

MCPHERSON, F. M. (1960). Conditioning and introversion–extraversion. Unpublished M.A. thesis, University of Aberdeen.

REID, R. L. (1964). Conditionability and eyeblink conditioning. *Proc. Annu. Conf. Brit. Psychol. Soc., Leicester*, 1964, *Bull. Brit. psychol. Soc.* **17**, no. 55, 16 A.

SYMON, S. M. (1958). An investigation of the relationship between conditioning and introversion–extraversion in normal subjects. Unpublished M.A. thesis, University of Aberdeen.

(*Manuscript received* 17 *June* 1965)

Brit. J. Psychol. (1965), 56, 4, *pp.* 485–502
Printed in Great Britain

PUBLICATIONS REVIEWED

A Model of the Brain. By J. Z. Young. Oxford: Clarendon Press. 1964. Pp. vii + 348. 35*s*.

In this reviewer's opinion, Professor Young's latest book is his most important yet. Expanded from the William Withering Lectures delivered in the University of Birmingham, it gives a step-by-step account of the distinguished work on the octopus and its brain which has engaged the attention of the author and his colleagues for the past fifteen years. But it is very much more than this. Prof. Young has developed a 'model' of brain activity which goes far beyond the special problems of the cephalopod nervous system and which might well provide a general basis for the explanation of behaviour in higher forms, not excluding man.

Basically, Prof. Young's model is a cybernetic one. By this is meant that we are once again asked to envisage the organism as a homeostat, to explore the analogy between brains and computers, to concern ourselves with the coding and processing of 'information', and to conceive of the higher nervous centres essentially as result-classifying systems. But Prof. Young is much too good a biologist to let cybernetics run away with him. His model is firmly rooted in the realities of nervous structure and checked wherever possible by the findings of experiment. As he himself puts it: 'The difficulties of building a realistic model . . . drive one at times to despair and to think of the brain as a "soup" of chemical processes or electrical fields. But then one's courage returns, one continues with the microscope and the microelectrode, and is drawn back to the ideal of Cajal and Sherrington, namely to understand the system in terms of the pathways and units that can be exactly studied in it'. Psychological model-builders: please note.

To readers of this *Journal*, the main interest of Prof. Young's book will probably lie in the extensive use made of the results of behaviour experiments. Several of the leading ideas stem directly from the long series of studies on discrimination learning in the octopus which he and his colleagues—several of whom are experimental psychologists—have carried out with noteworthy success over the past ten years. Indeed one gains the impression that, at the present stage of things at least, behaviour studies cohere much more closely with neuroanatomy than with neurophysiology. After all, if we accept that the brain is, in an important sense, the instrument of behaviour, may not its 'functions' be revealed more directly in behaviour itself than in the electrical or chemical properties of its component units? If so, have we not much to learn from studying, as Prof. Young has so elegantly done, the grosser functions of large regions of the brain in relation to definable (and within limits measurable) patterns of behaviour? This is not, of course, to undervalue the contribution of neurophysiology, which in the last resort will doubtless win the day. But it may well prove expedient, *pro tem*, for the experimental psychologist to link himself with the neuroanatomist and the neurosurgeon rather than with the expert wielder of the microelectrode.

If criticism is to be offered, one might perhaps object that Prof. Young has tried to go too far too fast. He himself warns the reader that at times he may feel he is learning too much about cephalopods and too little about brains. He may even feel that he is learning too little about cephalopods. After all, octopuses do other things besides discriminating shapes, and one is tempted to ask whether a model of the brain which leans so heavily on the existence of this capacity is equally applicable to all the other things which octopuses do. One may also wonder whether the simple Thorndikean principles which Prof. Young invokes to explain learning really meet the facts of the case. Will no sophisticated behaviour theorist come to the rescue and exchange his rat for a squid?

Seriously, though, this is a good book and a hard book. It really does deserve careful reading and will abundantly repay it. Professor Young combines scientific attainment of a very high order with an intellectual flair for which distinguished scientists are not always remarkable. Every psychologist worth his salt should read this book with the sort of concentration that he applied to reading in his undergraduate days—which the present reviewer was fortunately able to do as its appearance coincided with a period of sabbatical leave. He would strongly advise his colleagues to apply for sabbatical leave at the earliest opportunity.

O. L. Zangwill

Physiological Psychology. By C. T. MORGAN. 3rd ed. New York and London: McGraw-Hill. 1965. Pp. vii + 627. 72*s*.

This latest edition of what is recognized as a standard textbook on physiological psychology retains much of the basic groundplan of its predecessors but appears with a more attractive layout. The change of print and greatly improved figures make for a book which from the start gives the impression of greater liveliness than the earlier versions. The main changes which have taken place in content are a reduction in the number of pages dealing with sensation and motor function and an expansion of the coverage of motivation and learning.

The second edition appeared just before the large expansion of work in the field of physiological psychology and neurophysiology which took place in the 'fifties and, for instance, missed by a hairbreadth the 1949 paper of Moruzzi and Magoun which, at least partly due to the later advocacy of D. B. Lindsley, brought the reticular formation to the notice of psychologists. Another example at random of what could not be in the second edition which has now become very much part of the general knowledge in this field is the work of Olds and Milner. In this rapidly expanding field omissions such as these two examples are risks that any author has to take by choosing to publish at one time rather than another. The present edition has repaired these and many other omissions and includes the major developments since 1950.

The author states in his preface that, as with the earlier editions, his emphasis is on 'experimental work done with animal subjects rather than on human clinical studies'. In making any criticism of the book's coverage it is worth while remembering this deliberate limitation of field. It might however perhaps be reasonably suggested that there is a third element of physiological psychology that could deserve inclusion in this volume, namely non-clinical experimental work on human subjects, which is defined separately by some workers as psychophysiology. We thus find very limited mention of work on, for instance, measurement of cardiovascular variables, on the GSR, on the determinants of eye-closure, all of which it is important to know about if much of the present spate of work on the conditioning of these variables is to be well understood. It must however be recognized that the book is already a long one, and there is probably already scope for a companion volume which sets out to cover those aspects which Professor Morgan deliberately omits. Provided that the student recognizes that the contents of this volume are by no means all he needs for the study of physiological psychology for his first degree, he will find this a most useful work and probably easier to use than its well-established predecessors.

P. H. VENABLES

The Image and Appearance of the Human Body. By P. SCHILDER. New York: Wiley. 1964. Pp. 353. 13*s*.

The linking of scientific disciplines by relevant generalizations is potentially an illuminating but also hazardous undertaking. The hazard is illustrated by Dr Paul Schilder's attempt to unite neurology and psychoanalysis, now reprinted as a scientific 'classic'.

In its first part, this work deals with neurological syndromes supposedly implicating the 'body image'. No physiologically intelligible definition of this entity is attempted, and the miscellany of clinical observations quoted share little more than the common factor of some alteration in sensory acuity over some part of the body surface. Facile analogies with disordered psychodynamics, touched on in the first part, become rampant in the second. Dogmatic and often incoherent suggestions are thrown out without any serious attempt to marshal evidence or to support them by reasoned argument. The result is a series of intellectual firecrackers, too bizarre to persuade, too vague to test experimentally. It undervalues both neurology and psychoanalysis as topics for rational and consecutive exploration.

To create a neurology of psychoanalysis or indeed to elucidate the psychodynamics of neurological disease would be a valuable achievement. Anyone who makes the attempt might obtain some relevant insight from one of Schilder's stray remarks. But he will have to adopt a more sober and scientific approach if he wishes his contribution to be taken seriously.

ELIZABETH K. WARRINGTON

Patterns of Sexual Behaviour. By G. S. FORD and F. A. BEACH. London: Eyre and Spottiswoode. 1965. Pp. vi + 330. 30*s*.

This textbook, reset in its new edition but not rewritten, has probably contributed a very great deal to the perceptible change in social attitude of which it was itself a product. With the work of the late Prof. A. C. Kinsey, it belongs to the field of descriptive natural history—the continuation of the Havelock Ellis tradition. It owes little, on the one hand, to psychoanalysis, of which its authors seem remarkably innocent: on the other, it foreshadows, but does not really develop, the comprehension of human sexuality in terms of primate evolution which makes Freudian psychology immediately intelligible to biologists. Psychoanalysis and biology have thus joined hands, as it were, behind the back of this particular book—though with help from it, and from the discipline of field primatology to which Prof. Beach has contributed so much groundwork.

Its real importance—increased in this country by the fact that it appeared originally as a *Practitioner* text and reached the medical profession—has been in bringing home to a large number of readers, lay and professional, the true scope of human diversity and its continuity with that of animals. We cannot, it is true, base moral or clinical judgements on the behaviour of white rats and Trobriand islanders, but a knowledge of anthropology and natural history has helped to make untenable the stifling idea of sexual normality which handicapped both psychiatry and intelligent thinking about sexual behaviour in the early years of the century. We can judge the size of the advance by re-reading the reprinted works of Krafft-Ebing which are now being paperbacked—an untimely revival of ancient ghosts if ever there was one. Ford and Beach did not effect this very necessary change, but their book was one of the minor locomotives of history, if only because, like Kinsey, it got over to the reading public and even to the least educable of experts. There is a sense, as psychotherapists will know, in which the change is skin-deep only—the removal from our public posture of guilt over such matters as orogenital play, masturbation and homosexuality has not necessarily penetrated to the bottom of the individual's self-valuation, whatever he may say about them. But the gain is real, and at least we have stopped making things worse by writing books which foster such guilt, and have begun to come to terms with our own diversity. In an odd way, animals and 'savages' help us in doing this as toys and drawings help children—there may be no rational ground for feeling relieved that porcupines masturbate or that Socrates was openly bisexual, but at the irrational level it helps. One would wish this to extend beyond the sexual field—to the acceptance of our aggressions and death-wishes, for example—and this may well be the future trend of education in mental health.

So much has been achieved by this particular book that it seems ungenerous to wish that the authors had taken this opportunity of amplifying it. They may have been right in sticking strictly to their descriptive and non-evaluative posture, in confining their anthropology to tribal societies, and their primatology to observed physical mating-behaviour. In leaving out the behaviour of large civilizations and the phylogeny of sexuality as reflected in the primate family, they may have been resisting a temptation to theorize, but I, for one, wish they had been able to succumb to it, even at the cost of a thicker volume. The range of acceptable behaviours in complicated societies, more comparable to our own—Greece, Rome, the ancient Orient—was historically the source of the Englishman's first insights into the uncomfortable rigidity of his own sexual posture. It has a place in our comprehension of ourselves, but perhaps it needs another book by different authors. The only important thing which is missing from the primatological side of this study is a more explicit statement of the growth in the importance of sexual behaviour with the development of the primate family. This in its final human expression explains the part which sexuality plays in generating our own anxieties—and the fact that we had to select it as the first field for the reform of our own self-estimate. The omission is tantalizing, when even the religious are ceasing to insist that the main importance of sex in man is reproductive. But when all this is said, and when due allowance is made for the brisk anecdotal pace of some of the anthropology, Ford and Beach remains one of the more important contributions to mental hygiene in our century. If it appears widely on the bookstalls it will do much good and little harm, however cheerfully physical its assessment of such matters as premarital intercourse, which have their profound interpersonal implications. For the weaker brethren, correctives abound.

ALEX COMFORT

The Natural History of Aggression. Edited by J. D. Carthy and F. J. Ebling. London and New York: Academic Press. 1964. Pp. 159. 30*s*.

In this symposium volume a number of zoologists, psychologists, sociologists, anthropologists and historians try to find some common ground on which to discuss aggression. Needless to say, it is the zoologists who have the easiest task in the presentation of their material. They, at least, have a consistent definition of aggressive behaviour and they can relate it to the social and ecological background of their animals. Perhaps the most important single contribution in this volume is that of Hall who gives an excellent review of some of the new observations on monkeys and apes in the wild. It is striking how little overt aggression is seen, and there is often complete tolerance between groups, which may, for example, mingle at a common drinking-place. Lorenz provides an attractive account of the ritualization of fighting in a variety of animals, by which its signal value is enhanced and its potential for actual damage reduced.

There seems to be little agreement on how to define or delimit human aggression. Cannibalism, suicide, nail-biting, torture and nuclear warfare are all mentioned as aggressive activities. Hill, in his account of aggression and mental illness, Freeman, who approaches the subject from an anthropological viewpoint, and Storr, who writes on 'Possible substitutes for war', all consider that there is an inborn aggressive tendency in man. Burton, who deals with aggression in the setting of international diplomacy, doubts this. There is further confusion on possible links between the carnivorous proclivities of the Australopithecines and aggression. As Lorenz and others point out, most animal carnivores show well-controlled intraspecific aggressiveness. There can be little doubt that the development of tools by the Australopithecines resulted in a sudden and disastrous increase in what might otherwise have been minor scuffles. Man has not been able to evolve appropriate inhibitions against such attacks.

Whilst most of the contributions have points of interest, they reveal distressingly little real contact between those who study animals and those who study man. It would have helped to provide more emphasis on human aggressive behaviour at the individual level and some ways in which it can be altered by experience. It is along such lines that links with animal aggression may best emerge.

In conclusion, I must pay tribute to Laver's contribution entitled 'Costume as a means of social aggression'. To read Mr Laver on the hierarchical, utility and seduction principles in the history of clothing is both informative and completely fascinating.

AUBREY MANNING

Crime and Punishment in Britain. By N. D. Walker. University of Edinburgh Press. 1965. Pp. xiii + 367. 70*s*.

The greater part of this lively and well-written book is devoted to a description of the British penal system and a shrewd analysis of the manner in which it functions. The author begins by examining the scope and limits of the criminal law; he then proceeds to review the pattern of detected crime in contemporary Britain, and contributes some penetrating comments on the validity of published criminal statistics as indices of the real incidence of illegal behaviour. Later sections of the book deal with the aims and assumptions upon which present-day penal practice is based, the mechanism of the sentencing process, and the varieties of punishments and treatments that are available to the courts. The last part is concerned with the characteristics of three special classes of offenders (the psychiatrically abnormal, females, and recidivists), and considers the ways in which they may be dealt with. Dr Walker has a keen eye for the essentials of an administrative system (sharpened, no doubt, during his distinguished career as a civil servant); he guides the reader deftly through the intricacies of statute and judicial precedent, and provides a clear and useful account of contemporary penal methods. There is a scholarly treatment of the ethical problems of responsibility, blame and punishment. In short, this is a very good introductory textbook of penology.

It is a pity that Dr Walker decided to include a discussion of contemporary explanations of the origins of criminality. This is, of course, a conventional ingredient of criminological texts; the difficulty is that the theoretical contexts in which these explanations were conceived are so diverse that a brief review in non-technical language is bound to be hopelessly superficial. The author's skill as an expositor does not help him here; indeed, it makes matters worse, for each

view is stated with such neatness and lucidity as to give the impression that the issues are simple and clear-cut, which they are not. For instance, a genuine analysis of the points of correspondence and divergence between psychoanalytic and Eysenckian accounts of criminality would surely constitute a formidable intellectual feat, but the reader is given no inkling of this; he is told that the disagreement between these schools of thought may be 'reduced to a question which might conceivably be answered by empirical investigation'. This brisk and summary way of dealing with complex issues is characteristic of Part II of the book ('Explaining and Predicting Crime'), and leads to some curious pronouncements. It is startling, for example, to read that 'since the Intelligence Quotient is a measure of the relationship between a person's chronological age and his mental age, it is in theory legitimate to say that a boy of 12 with an I.Q. of 75 (i.e. a mental age of 9) is as intelligent as a 16-year-old with a mental age of 12'. It is odd to find the term 'paranoiac' still being used to describe someone suffering from paranoid schizophrenia, although this might be defended on the ground that the word is familiar to laymen; the same cannot be said for 'ataractic', which appears without explanation later in the book.

Dr Walker has some interesting things to say about the strategy of research; the treatment of prediction techniques is a model of lucidity. His view of methodology is occasionally a little old-fashioned. Thus in spite of Popper and Toulmin he asserts that an hypothesis can be verified, but only by forecasting the outcome of observations which have not yet been essayed; 'pseudo-predictions', in which it is demonstrated that known facts are necessarily implied by a theory, are said to support but not to test it, 'since this knowledge may well have influenced the form of the hypothesis'. This suggests a rather unsophisticated notion of the relations between theoretical terms and operational variables.

But these are comparatively minor defects in a stimulating and most readable book which—despite a generous sprinkling of misprints—is good value even at seventy shillings.

GORDON TRASLER

Ingratiation: a Social Psychological Analysis. By EDWARD E. JONES. New York: Appleton-Century-Crofts. 1964. Pp. xi + 211. No price given.

Professor Jones has, in the past, experimented extensively in the area of person perception, contributing notably by his work on the ways in which interaction, and consequent judgement, are affected by the orientations and expectations of the participants. The present study is a natural extension of that work. He explains that the study of ingratiation sprang from the demonstration that our judgements of others are strongly affected by the face they present to us. 'Ingratiation' refers to presentations which are motivated by the desire to create an attractive face. It further implies, by definition, that the ingratiator is perceiving the interaction situation in a different way from the 'target' person while giving out signals intended to assure him that they are sharing the same perceptions.

In a conceptual analysis of ingratiating behaviour, Professor Jones lists possible strategies and tactics, the main tactics distinguished (these are to become experimental paradigms) being (i) the use of compliments to enhance the self-esteem of the target person, which may have to be exercised in a devious way since the ingratiator's goal may not be achieved unless his motives remain concealed; (ii) conformity to the opinions and behaviour of the target; (iii) self-presentation in a manner designed to increase one's attractiveness. The implications of various situational factors, such as differences in the power relations of the participants, applied to these tactical manœuvres, give rise to a number of hypotheses, tested by a programme of experiments.

The work of conceptual analysis and hypothesis formation draws upon other theories in social psychology—Festinger's cognitive dissonance theory, Heider's balance theory, Thibaut and Kelly's analysis of power relations and Goffman's ideas on 'face-work'—and is set within the overall framework of investments, rewards and costs, as used by Homans, among others. Thus, apart from the gains of the experimental programme, this work also contributes to the integration of theory in the areas of social interaction and person perception. However, the yield of the experimental harvest is a little disappointing. Several hypotheses failed to achieve confirmation, and all too often one suspects that the fault lies in shortcomings in the design and testing-out of experimental conditions. Consequently, confidence even in positive results is undermined. Above all, the experimental work lacks replication, despite the fact that the eight or nine experiments collated here have all been published independently. The main value of the book

therefore lies in its theoretical analysis, and this could have been increased by further work on the supporting experiments. The concluding chapter shows that the author was by no means unaware of the inconclusiveness of the study. He says himself that it would be useful to verify and replicate the relationships reported here, although he appears finally dubious of the value of laboratory experiments in the analysis of phenomena of this kind. Ingratiation phenomena, he says, studied by such a strategy, often turn into something else, so that we end up with a 'vaguely bound cluster of social influence strategies each having its own phenotypic properties'. However, it can surely be argued that the demolition of the categories of 'common sense' is never a useless activity in psychology.

THELMA VENESS

Social Psychology: The Study of Human Interaction. By T. M. NEWCOMB, R. H. TURNER, P. E. CONVERSE. New York and London: Holt, Rinehart and Winston. 1965. Pp. xii + 591. $8.50.

Newcomb, Turner and Converse have written a textbook on social psychology that has several major assets: it presents a coherent account of the subject, which in itself is no minor achievement; it is firmly based on empirical work, containing some sixty detailed research illustrations and using fully the senior author's recent work on the process of becoming acquainted; it aims to combine psychological and sociological approaches; it does not make a shibboleth out of any particular theory; and it is written throughout in readable English.

The text is organized in four parts. Part I is entitled Individuals' Attitudes; Part II, Processes of Interaction; Part III, Group Structures and Properties; Part IV, Interaction in Group Settings. There is an excellent appendix on attitude measurement, a competent one on survey research and the measurement of public opinion, and one on Bales's interaction process analysis which, the authors claim, is 'one of the most widely used methods of describing interaction among members of small groups', a claim which their own research illustrations do not bear out. A similar situation occurs with regard to Guttman scaling which is explained with exemplary clarity in the appendix, but not mentioned anywhere else in the book, unless one takes the statement on page 51 on the need for studying multidimensional stimuli in their full complexity as an implied critique of Guttman scaling. We are all wont to say that major problems can be tackled only to the extent that appropriate methods are invented. True enough; but methods which were eminently suitable for past problems in social psychology have a way of lingering on in textbooks as if they were still suitable for the more complex tasks of current social psychology.

In the text itself the authors treat attitude as the core concept of social psychology, referring throughout to the individual as being properly represented in social psychology by the organization of his attitudes. The early exposition of attitudes forms the coherent thread which goes through the entire book, even more so than the authors' professed major theme, which is the concept of interaction.

Since there are now several good textbooks on social psychology available, it becomes necessary to look at each newcomer in a comparative fashion. This book is a great improvement on Newcomb's *Social Psychology* published in 1950. Not only because it is inevitably more up to date; of the almost 300 references, about 190 refer to work done after 1950; it is also more adult and sophisticated in its tone, treating the undergraduate students as intelligent young adults should be treated. However, the very assets of the book imply also some of its shortcomings. As the authors recognize, the aim of producing a coherent text compelled them to omit several areas of work in social psychology which were covered in Newcomb's earlier volume, and are covered in Krech and Crutchfield. The terms 'personality', 'culture' and 'socialization' are not even mentioned in the index. If this can be justified, a more regrettable omission is Roger Barker's ecological work, which has greatly developed in the last fifteen years and presents one of the few new approaches in social psychology which deal adequately with the psychological meaning of both physical and social environment. Sociologists may feel that their work is under-represented.

As a result of such omissions the volume will have to be used in conjunction with other texts and selectively, rather than as *the* basis for a course.

MARIE JAHODA

Parental Behavior and Sources of Information in Different Social Groups. By T. Nummenmaa and M. Takala. Jyvaskyla Studies in Education, Psychology and Social Research, 11. 1965. P. 52. No price given.

This monograph presents the results of a survey of the maternal attitudes and child-rearing practices of 785 Finnish mothers. Using school records, the investigators chose a nation-wide sample of families, with the restriction that every family should contain at least one child aged 3–4 and one aged 10–11 years. This extensive sample was stratified, so as to make possible a comparison between mothers in different social classes and in rural and urban districts. The mothers were interviewed at home using a 140-item questionnaire; questions covered health, physical development, toilet training, sex education, expression of affection, educational achievement and plans, the child's frankness with the mother, religious behaviour and punishments. It is perhaps a pity that a few of the items referred to events which had taken place some time ago—mothers' memories about toilet training, for example, are not always to be taken at face value—but on the whole the questions were sensible and straightforward.

The interviewers were twenty students of psychology of whom more than half were women. They were given a ten-day course in interviewing techniques, and undertook the field-work in their own home districts. The authors write: 'The fact that the number of interviewers was rather high, and the fact that each of them interviewed representatives of each social status probably safeguards against any systematic effects due to interviewers.' There are, however, certain dangers inherent in trying to combine a research-training exercise with a serious research project, the most obvious being the temptation to include the results from all the interviewers, good, bad and indifferent, and to treat them all with equal respect. Possibly the investigators were aware of such difficulties, and this would provide an explanation for the curious way in which they decided to analyse and present the results of the study. Although most of the original questions seem to be given in full, it is not possible from the report as published directly to compare mothers in the eight different groups (4 social classes × 2 types of district) according to the answers given to specific questions.

Instead, the individual questions have been grouped together in small clusters, defined partly by area of content and partly on the ground that the answers to the questions within each cluster are positively correlated for each of the eight groups. We are not told the degree of correlation thought necessary, and indeed it seems that even the basic requirement of positive correlation is waived in a 'few exceptions'. The authors then present only means of cluster-scores for each of the sub-samples. From the practical point of view, this method of presenting the first-order results is particularly frustrating because it is quite impossible for the reader to apply orthodox statistical tests to the results for any given cluster in which he may be interested. This is especially irksome where the amalgamation of results seems unduly arbitrary—where, for example, it is impossible to distinguish between incidence of corporal punishment in the upbringings of mother, of father, and of the present children's playmates. To confuse the issue still further, the report then plunges into a factorial analysis based on some, but not all, of the cluster-scores; and the final interpretation rests upon six factorial dimensions labelled 'general information', 'candour', 'religious behaviour', 'personal communication', 'schooling' and 'professional advice'.

With such a tortuous statistical approach, it is not surprising that the original aim of this investigation—'to describe what sources of information are used by parents in various social groups in connexion with different problems encountered in child-rearing and upbringing, and what procedures are eventually adopted'—is almost lost sight of. An attempt in the final summary to tie up intentions with results is not convincing. To quote one conclusion: 'In general: when the parents have to cope with an "institution" (of these the school, the church and medical science were dealt with in the present study, but the results could well apply to economic and governmental institutions), parental behaviour would be affected by the opinions of experts, though not in the same way or by the same sort of experts in all cases.' This statement, vague as it is, is only minimally substantiated by the evidence published here; and to read that the authors think they have 'dealt with' the mothers' relations with 'the school, the church and medical science' makes one search through the interview questions for missing pages. Indeed, since the only question relating specific advice to specific action, even by implication, is that concerning drugs prescribed for the child, it is difficult to see how very definite conclusions could

be reached on this topic. Here again, however, one must concede the limitations set by the use of student interviewers.

The sad thing is that there was obviously a great deal of interesting material in the original results of this survey. In Finland, as in this country, middle-class parents are more likely than working-class ones to demand early toilet training, more likely to encourage physical expressions of affection, to be frank about where babies come from, to read and to possess books, and to take an active interest in their children's education. The sources from which these attitudes are derived are many and complex; but they are not entirely unamenable to investigation. It can at least be said that this inquiry has cleared the ground for a more direct and probing attack upon a subject which must be immediately relevant to any study of the child in his family.

ELIZABETH NEWSON

The Family and Individual Development. By D. W. WINNICOTT. London: Tavistock Publications. 1965. Pp. 181. 30*s*.

'It is better to exist in prison than to become annihilated in meaningless compliance' (p. 151). 'We are poor indeed if we are only sane' (p. 61). These quotations highlight a cornerstone of Winnicott's philosophy: social adjustment at the expense of individual integrity is scarcely worth having. Not only do symptoms have utility (as when he describes thumb-sucking and nail-biting as 'a sort of built-in psychotherapy') but mental illness itself, including delinquency and psychosis, may represent an individual's only hope of breaking down the façade of a 'false self' and discovering his true identity. This radical view of psychopathology, as yet little appreciated by the general public or the caretaking professions to whom these lectures are addressed, opens up a new and arresting perspective on many current problems.

How can one tell a 'true' from a 'false' self? This point might have been more explicitly developed, but we are given certain clues. There is recurrent emphasis on 'that spontaneity which alone makes life worth living' and on 'the creative impulse, that which (as nothing else) proves to the child that he or she is alive'. Creative spontaneity is made possible by adequate and reasonably strict mothering, which provides an example of moderation for the taming of the primitive superego. In chapter 2 we learn that 'if the mothering is not good enough, then the infant becomes a collection of reactions to impingement'. The inference is that such fragmentary, more or less stereotyped reactions cannot integrate into a whole person; only reliable early relationships can build up a coherent individuality that will be true to the inherited potential.

In adolescence, the integrity of the self is severely tested. Winnicott welcomes the current exacerbation of 'adolescent doldrums' as an inevitable concomitant of the disappearance of military discipline and preparations for war; does it not also apply then to the youth of those countries which are still preparing for or engaged in war? He insists on the rightness of the adolescent's need to feel real, and to test his reality for himself. This entails the rejection of all compromises, including those involved in adult moderation, and the need to prod society into reacting with antagonism which can be met with antagonism. Society, he adds, should 'meet the challenge' (how?), but not try to cure the doldrums, which are essentially healthy.

Is this asking more tolerance than the community can give? What in fact should be done with hooligans whose reality-testing, however valuable to themselves and their less disturbed admirers, jeopardizes the safety of others? This question is not answered directly. But since the author believes that a sense of deprivation always underlies antisocial behaviour, a partial answer can be found in two chapters dealing with those deprived and maladjusted children for whom residential care is found necessary.

Despite his profound individualism Winnicott is no anarchist. Nor does his deep faith in life ever fall into sentimentality. There is always the saving, wry remark, or the sober reflexion on a case that went wrong; or a shrewd rendering unto Caesar with regard to practical management, as in the recognition that delinquents whose capacity for personal response has been stultified early will always need firm discipline.

It is tantalizing at times that the lecture form forces him into *ex cathedra* statements, insufficiently supported by evidence, on some striking and controversial point which might have been better substantiated had space permitted. But the form has compensations in conciseness of argument, informality and appropriateness for the non-psychiatric reader; while the original and challenging quality of the thought makes this book most rewarding for professional and layman alike.

TERENCE MOORE

Organisational Stress: Studies in Role Conflict and Ambiguity. By R. L. KAHN and others. New York: John Wiley and Sons. 1964. Pp. 470. 70*s*.

Over the past decade there has been a spate of writing by applied social scientists on the impact of the organization upon the individual. The focus of attention on organizational stress is more than justified in the prevailing conditions, since, by tradition, social institutions have restricted the individual's growth as much as they have helped in meeting his deficiencies. Here the problem is studied within the context of role theory and the findings clearly demonstrate the contributions to the individual's work-related tensions of conflicting demands by associates and ambiguity of work roles.

This book is the result of five years' research at the University of Michigan supported by the National Institute of Mental Health. The work was carried out on two fronts: an intensive study of 53 'focal persons' and members of their 'role set' in six firms (role set is the term used to describe all those people who have expectations of the focal person) and a survey of a national sample of 725 working people in the United States. The principal interests of the authors appear to have been to explore the extent of job tensions, measured by attitude scale, and to link this with role conflict and role ambiguity in an attempt to identify the interactions with organizational processes, interpersonal relations and personality variables.

The major causes and effects of stress are covered in some detail with the arguments supplemented by case studies and statistical analyses. Though the reader is apt to find himself at sea in the wealth of material, the book is well structured. Sound use of headings, carefully outlined aims and definitions, consideration of results within the framework of a role dynamics model and the use of short chapter headings, all contribute to its readability. Only in the description of the method, which is not easy to follow, and in the conclusions, which the authors begin by presenting new material of a summarizing kind, does this work fail to maintain its high standards.

The results from the national sample suggest that only one job in six is free from stress. The studies show that conflicts in the expectations of 'role senders' are positively associated with subjective experience of tension and that these conflicts are mostly of a hierarchical kind. Role ambiguity is shown to be one of the results of technological factors, in particular technical change, and is also linked with higher job tensions. It would be interesting to duplicate this American study in Britain. Kahn and his associates note that work-overload is a major source of stress among American workers: one suspects that lack of responsibility and underemployment may be more important contributors here.

In their attempts to present a comprehensive picture, the authors have tended to develop notions which run beyond the bounds of their data. Some confusion is also caused by the authors' failure to define their concept of mental health. Repeated reference is made to effective performance, a notion we would expect to find in any discussion within the context of positive mental health, yet there is no empirical evidence on that subject. Although the descriptive part of the exercise seems highly feasible, the practitioner is left to solve his problems of stress, organizational change and the like by first identifying an entity, the 'focal person's role set', which is larger than the individual and yet less precise than the group. Who knows? The results may be worth the effort.

This is a good textbook for the social scientist who is interested in organizational problems and role theory and it may well prove useful to the well-informed practitioner in search of a model within which to consider the shortcomings of his firm.

DENIS PYM

Social Psychology in Treating Mental Illness: an Experimental Approach. Edited by GEORGE W. FAIRWEATHER (ed.). New York: Wiley. 1964. Pp. xii + 300. 60*s*.

This is another study of an attempt to tackle the problem of rehabilitating chronic mental hospital patients. Although a number of such studies have been published in recent years, further studies are welcome if they can be shown to provide new data. This book does provide new data. It has been edited by a clinical psychologist and is 'the culmination of three years of work in a continuing social psychology and psychiatric research effort aimed at finding a realistic and pragmatic solution to the problem of the chronically hospitalized mental patient'.

Twelve authors have contributed, but apart from the editor it is not clear whether they are psychologists, psychiatrists, social workers or nurses.

The data have been well organized and given detailed statistical treatment. At the same time an attempt has been made to present the material in non-technical language, so that administrators and others not specifically trained in technical research can understand the conclusions and the data on which they are based.

The study begins with the observation that the protective hospital setting with its minimal demands on the patient cannot be expected to prepare him for the more rigorous demands of community life outside. An attempt to bridge this gap is made by establishing groups of patients within the hospital who would return as units to the community. The behaviour before and after discharge of patient-led task-groups of this sort was compared with that of other groups organized in the traditional manner of superordinate staff and subordinate patient roles.

As might be expected, and as was found in the Claybury study, the effect of changing some parts of the hospital social system was felt all over the hospital. Three chapters deal with this interesting problem.

Although there were clear advantages in change of behaviour of the small task-groups compared to the traditional ones, these gains were not always carried over into community life after discharge. For example, although significant differences were found in the small task-groups after discharge compared with the traditional groups in employment, friendships and verbal communication with others, no differences were found in the rate of rehospitalization, appraisal of illness (i.e. symptoms complained of), drinking behaviour, nature of residence, membership of community groups and involvement of leisure-time activities.

The authors conclude that training for community living should be done in the community, and that organized patient-led task-groups of mental patients as a rehabilitation device should be carried out in a community setting rather than in hospital.

Anyone wishing to work in this important area should read this book with care before planning his research.

RALPH HETHERINGTON

Case Studies in Behavior Modification. Edited by LEONARD P. ULLMANN and LEONARD KRASNER. New York: Holt, Rinehart and Winston. 1965. Pp. 401. 64*s*.

This book contains 50 case-studies, mostly reprinted from a variety of journals and concerned in the majority of cases with deviant behaviour in children and schizophrenics. The stress throughout is very much on operant conditioning and practically all the work included in this book is American. There is also an extended introduction, running to 67 pages, entitled 'What is behaviour modification?'; this discusses the nature of mental illness and the value of current nosologies, the bases of the practice of psychotherapy and the proper role of the psychotherapist. Also provided are introductions for each of the five sections into which the material is divided: severely disturbed behaviours, classical neurotic behaviours, deviant adult behaviours, deviant behaviours in children and mental deficiency.

The editors point out that their book differs in many ways from *Behaviour Therapy and the Neuroses*. None of the articles appearing in these anthologies overlap and the types of cases studied differ very much from one volume to the other, some 70 % of the material in the new book being devoted to children and schizophrenics as against less than 20 % in the earlier volume. They say that 'in great measure the availability of this much material illustrates the viability of the field. Only six of the fifty articles printed in the present volume were published before 1960, the date of Eysenck's *Behaviour Therapy and the Neuroses*.' They also stress that they have written for a more application-oriented audience, 'a matter which may reflect the increasing acceptance of behaviour therapy and behaviour modification'. The readership aimed at is both on the undergraduate, non-specialized level, and also among practising psychotherapists. The use of case-histories in this connexion derives from the desire of the editors to 'present material at an inductive level'. They use these cases as 'illustrations of what is being done to mitigate and change maladaptive behaviour', fully aware of the fact that such material is 'not evidence or validation in itself'. From this point of view this is an excellent production which should be in every psychological library, and familiar to every clinical psychologist. The choice of cases is first-rate, the production of the book impeccable, the price reasonable; and the intro-

duction and comments by the editors combine an easy readable style with a profound understanding of the issues involved, and a thorough theoretical and practical acquaintance with the whole field. Few readers unfamiliar with the subject matter of this book will be able to avoid experiencing a mounting feeling of excitement as they become aware of the great strides which have been taken in clinical psychology by experimentalists applying the laws of learning theory to patterns of abnormal behaviour. A field that once used to be a prison of spurious orthodoxies has suddenly been thrown wide open by groups of young rebels, many working in isolation and quite independently of each other; more new ideas have been generated in the last six years than in the 60 years preceding the publication of Wolpe's book which might be said to have formed the *fons et origo* of this movement, and there is little doubt that now that the log jam of ideas has finally been broken many great developments will follow. Books such as this play an invaluable part in making these new ideas available to students and others who may in due course be called upon to contribute to these developments.

H. J. EYSENCK

Can Psychopathology be Measured? Edited by HAROLD WHIPPLE. *Annals of the New York Academy of Sciences*, **105**, no. 15, pp. 813–926. $4.00.

The odd title of this booklet is matched by the contents, most of which are not concerned with the measurement of psychopathology, in the usual sense. The series of papers are preceded and succeeded by an introduction and postscript by J. Zubin which are worth reading, for, although the points made are not new, they are made very well. The first four papers describe research on schizophrenic patients, three of them comparing their results with normal control subjects. Pupillary reactions, the effect on speech of delayed auditory feedback, the use of the 'cloze procedure' for analysis of speech and the characteristics of schizophrenic speech. The quality of these papers is high and they will certainly suggest lines for further investigation. The paper by Burdock and Hardesty gives an account of the construction of a diagnostic inventory for children's behaviour. It is a bold attempt to tackle a very difficult subject, and such a scale is greatly needed. The last paper is a straight clinical paper on mental illness in the aged.

These papers are good and will certainly be quoted in the literature in due course, but it is difficult to see what links them together; certainly not the title of this particular section of the Annals of the New York Academy of Sciences.

MAX HAMILTON

Can We Explain the Poltergeist? By A. R. G. OWEN. New York: Garrett Publications. 1964. Pp. 436. $8.50.

The answer, it appears, is no. We cannot explain nor can we explain away the poltergeist. As a start, however, we can make certain tentative generalizations about the phenomena involved, and in this scholarly and scrupulous survey Dr Owen (mathematician, lecturer in genetics, Fellow of Trinity College, Cambridge) has gone about as far as it is possible in present circumstances to put the problem into perspective.

What makes his task such a ticklish one is that he has constantly to worry about the authenticity of the evidence that he is discussing. He rightly points out, however, that even one really well attested case may throw a flood of light on the more dubious cases. Fortunately he did find one case of recent occurrence where, after interviewing the witnesses and going into every detail, he was able to satisfy himself that none of the usual explanations in terms of suggestion, malobservation, trickery, earth-tremors, etc., was applicable. This was the Sauchie Poltergeist, where the disturbances (movements of furniture, loud raps, etc.) centred around an eleven-year-old girl whose parents had just previously migrated to Scotland from rural Ireland. The age, here, is fairly typical, and as to sex we find that girls become poltergeist foci roughly twice as frequently as boys.

After studying the author's account and having myself met one of the witnesses (he was the family doctor) I am quite sure that Owen was justified in committing himself to a positive conclusion in this instance. What makes the Sauchie case exceptional is that the witnesses include five professional persons of good standing (a clergyman, three qualified physicians and a primary school teacher). Of course, how far such a case justifies us in accepting some of the more fantastic

cases cited in this book will be a matter of opinion. As Owen points out, in this field credibility tends to become a very relative concept.

From a purely psychological point of view, poltergeistery does make a certain sort of sense. It would seem to be one way, if a very unusual one, of coping with repressed anxieties: Jung once called it an exteriorized complex. From the physical point of view, on the other hand, the problem of trying to make intelligible the *modus operandi* is truly staggering. The trouble here is that we are not just dealing with some indiscriminate explosive force of unknown origin but with seemingly controlled movements of objects that create the effect of an invisible imp playing pranks. It would be wrong, however, to make our bafflement an excuse for ignoring a phenomenon which, though rare enough in absolute terms, is of fairly regular incidence. (Owen reckons that a case gets reported from somewhere about once a year on an average). Let us hope that, as a result of Dr Owen's invaluable contribution to the topic, men of science will be more disposed in future to take an interest in these matters. It is only a pity that its usefulness as a handbook is much reduced by the absence of an index.

J. BELOFF

Ergonomics—Man in His Working Environment. By K. F. H. MURRELL. London: Chapman and Hall. 1965. Pp. 496. 63*s*.

For some time now there has been a need for a textbook of ergonomics. With the exception of brief reviews such as Fitts's chapter on engineering psychology in Stevens's handbook and the contributions by Fitts and by Taylor to Koch's *Psychology: a Study of a Science*, vol. 5, the student is faced with an unhappy choice between handbooks on equipment design which consist of specific recommendations, usually presented in a comic-strip format, and original research reports.

One problem in writing a text is that the subject is of interest to workers with widely varied backgrounds and it is difficult to cater for all in a single volume. Murrell has opted to write primarily for engineers and managers, and in order to give them an understanding of the subject he has prefaced the main section of the book with 125 pages of background material on human biology.

Human biology is covered by chapters on Bones, Joints and Muscles as Mechanical Devices; Metabolism and Heat Regulation; Body Size; The Nervous System; and, finally, Man as a System Component. This course is generally well planned and it brings together much information which cannot elsewhere be found in a single source. The ergonomist will particularly appreciate the critical review of anthropometric data and their current inadequacies. The presentation, however, is scrappy and uneven: some points are treated superficially while others are discussed in unnecessary detail. The writing is sloppy in places and references are scattered at random, so that, while much is straight assertion, some facts are carefully documented. Where references would be particularly useful, namely in providing guidance to further reading, they are absent.

Some caution should be given to those who will rush out to apply the smattering of knowledge they are likely to acquire by reading this section. Instead they are encouraged. Thus, after a brief discussion of referred pain in chapter v, the reader is given this advice: '. . . it must be remembered that the location of pain may be somewhat remote from its actual cause. But with this proviso in mind, it should be possible to make a reasonably intelligent diagnosis of the probable cause of the pain and make the necessary alterations in the equipment.' I think not.

The main part of the book is in three sections: (1) Design Factors, which is divided into chapters on Equipment Layout, Seating, Displays, Controls and Compatibility; (2) Environmental Factors, with chapters on Heat, Noise, Light and Vibration; and (3) Organizational Factors, in which he discusses Techniques of Measuring Efficiency, the Organization of Work (Rest Pauses, Pacing, etc.), Inspection, Shift Working, and Age.

Murrell has taken a rather narrow definition of his subject-matter and within his chosen field he has confined himself almost entirely to the factory and office work situations. His account of current knowledge in this area is well organized and clear. The chapter on the design of instrument displays is particularly good and shows a more critical approach than some writers adopt, quite properly emphasizing the need to examine carefully the purpose of a display before attempting to optimize its design. Throughout the book he shows an awareness of practical problems which will be appreciated by the engineer. There is some ambivalence between two

forms of presentation: some chapters read like research reviews, whereas others list recommendations. The review chapters lack the thoroughness to be expected of genuine research reviews. To give one example of this, the work of Colquhoun has been completely overlooked in the chapter on Inspection. This particular omission could have arisen because Murrell decided to present only field-studies, but with other topics laboratory experimentation has been described at some length.

The book is reasonably up to date, the cut-off year for references being 1963. There is no mention, however, of a number of recent trends such as work on the design of consumer goods, on devices such as telephones which have to be worked by the public, on the problem of writing intelligible instructions, on the relevance of verbal learning research to teaching codes such as those that postmen will put on letters to permit automatic sorting. Indeed the whole subject of training has been overlooked. The psychologist who buys this book to find out what ergonomics is about will receive the false impression that it is restricted to studies of tasks undertaken on the factory floor, and this is unfortunate.

Production appears to have been rather hurried. In chapter I there are some careless mistakes. The plates would have benefited from more care: some, such as photographs of noise-level meters, are unnecessary, some are inserted in the wrong place, while others are of appallingly poor quality.

H. C. A. DALE

An Exploratory Study of the Perception of Causality. By LAMBROS HOUSSIADAS. Monograph Supplement XXXVI of the *British Journal of Psychology*. London: Cambridge University Press. 1964. 22*s*. 6*d*.

It is a pity that the Monograph Supplements of the *B.J.P.* are not more widely known and supported. They afford, as near as can be, the ideal medium for the publication of research reports. The traditions and conventions of the universities make theses as presented for higher degrees almost invariably unpublishable. Even the full-length article in a journal is too cramping. The Monograph Supplement is just about right. It requires some pruning of the thesis, but enables the author to report his facts in sufficient detail and to develop his interpretation of the facts. Of this, the latest of these monographs is a good example.

Dr Houssiadas here reports a series of experimental studies which advance in important ways our knowledge of the phenomenology of the perception of causality; and, while following Michotte in his techniques, indicates some respects in which Michotte's interpretations of the facts require correction.

The experiments reported demonstrate conditions under which one object is perceived as changing the form of another object, compressing it or displaying its elastic properties. In another series of experiments the phenomenon of 'pushing with contact' is shown to occur when object A and object B move simultaneously.

Michotte, while opposing Hume on some fundamental issues, had accepted Hume's thesis that a cause-event must precede an effect-event. Houssiadas reports observations in which the two events can *appear* to be simultaneous, as when, in ordinary experience, the movement of a train seems to be simultaneous with but is judged to be the cause of the movement of its shadow.

Incidentally, but importantly, facts are reported which bear on discussions of causal action at a distance, on the distinction between 'agent' and 'patient', and on the relation of the concept of 'causality' to the concept of a 'thing' or 'continuant'.

The facts set out are important, and of interest in their own right to psychologists. They are not without importance to the many philosophers who today are interested in the analysis of psychological concepts and the validation of psychological findings. Michotte was perhaps a little over-concerned with the relevance of his findings to the doctrine of causality defended by Hume and his philosophical epigoni. Today, however, few authorities would deny that Hume tried to support an over-simple analysis of causal statements by an over-simple account of the psychology of perception. A distinction must be drawn between epistemological and psychological questions. Philosophers are concerned mainly with the question: how do we *know* that one event is the cause of another? Psychologists are concerned with such questions as: how can it *appear* that one event is the cause of another, even when we *know* that there is no such causal connexion? Hume would not have been—or should not have been—in the least disturbed by

Michotte's findings. He should have been delighted by the evidence that we make causal judgements which are demonstrably illusory.

On the analysis of the concept of 'cause' it is of interest to note that there is nothing in these experimental phenomenological studies to suggest that anyone ever gets the impression of there being a 'pure and simple' relation of causation of the kind which some writers seem to suggest. The impressions of causal actions which are perceived are impressions of certain specific kinds of action such as pushing, pulling, repelling, attracting, deflecting, intercepting, deforming, etc.

On the question of validation of causal statements it might seem odd that in the more striking cases in which there is an immediate perceptual impression of causal action *real* causal action is demonstrably absent. The psychologist *knows* when A really causes B, and *how* A causes B. He knows when and how A only seems to cause B. He must know this in order to design his experiments. But he does not know how A causes B *qua* psychologist but in virtue of his 'general science'. He knows how the card behind the screen rotates and how the variables of speed of rotation, etc., really produce the illusion of causality. He *infers* real causal connexions in accordance with the criteria of the natural sciences, which criteria do not include the perceptual impression of causality. The 'impression of causal action' is never used as evidence for real causal action.

There are indeed phenomenological differences between the case in which we judge that the horse is pulling the cart and that in which we judge that the cart is pushing the horse, but the validation of both judgements depends on further evidence.

The most difficult issue in these phenomenological studies is that of distinguishing 'immediate impressions' in perception and impressions based on past experience. The methodology of establishing such a distinction is frightfully complicated. Clearly, the judgement that the movement of a train is a cause of the movement of 'its' shadow while the movement of a shadow does not cause the movement of 'its' train is not based on watching innumerable examples of shadows 'following' trains or of trains following shadows; but it does depend, surely, on past experience which enables us to say what it is to be a train, and what it is to be a shadow. So, too, impressions of apparent causal actions displayed by shadows in a shadow picture on a screen are corrected by knowledge of how shadow pictures are produced. And so, *mutatis mutandis*, with the effects produced through Michotte's very ingenious instruments.

There are clearly several very different sorts of issues to be disentangled. In this monograph they are not completely disentangled—nor could this be reasonably expected. The report does, however, contain all or most of the facts required by psychologists and philosophers who may be disposed to disentangle them.

C. A. MACE

Mathematics and Psychology. Edited by GEORGE MILLER. New York and London: Wiley. 1964. Pp. 295. 42*s*.

Those who are not themselves mathematical psychologists have at best ambivalent feelings towards those who are. Much mathematical psychology seems difficult and inaccessible to the outsider. Moreover, there is the suspicion that a poor theory can be made to seem impressive if the mathematics are sufficiently elaborate. At the same time the growing need for mathematics is, one hopes, generally recognized. This book provides a first-class conducted tour of mathematical psychology and should increase understanding of what has been going on in this area. It consists of excerpts from papers with introduction and comment from George Miller. With one or two exceptions, the excerpts have been so chosen that they can be understood by those with only a limited knowledge of mathematics. Even where the mathematics may defeat the reader, Miller's lucid exposition of the author's ideas will be found highly illuminating.

Miller distinguishes *discursive, normative, functional* and *structural* applications of mathematics in psychology (but without implying that they can be sharply differentiated). A *discursive* application does not actually use mathematical reasoning: it merely uses mathematical notation as a convenient extension of natural language. Predictably, he gives an excerpt from Lewin as an example of this use of mathematics. He also includes a passage from Herbart on concepts as forces. A *normative* application involves the use of mathematics to discover the most efficient way to achieve specified goals or to formulate rules which we do or should use to guide our choices. This is a relatively novel use of mathematics to psychologists and all the illustrative excerpts are from the writings of non-psychologists—D. Bernoulli, W. S. Jevons, F. P. Ramsey

and J. Marshak. (The one from Marshak is from his review of games theory.) He does, however, describe how Ramsey's proposals for measuring utility and subjective probability separately were utilized by Davidson, Suppes and Siegal in 1957. A *functional* application of mathematics is the one which is most familiar and involves the statement of laws in terms of functional equations. Functional applications can be subdivided into determinate and statistical. Passages from Fechner, S. S. Stevens, Hecht and Hull are included as illustrating the former. Among his examples of statistical applications are excerpts from Thurstone, Estes and Tanner, Swets and Birdsall. He also includes part of his own paper 'What is information measurement?' and an excellent passage from Bush and Mostellar on the value of stochastic models. Finally, *structural* applications involve 'the use of mathematical notation to represent the reduction in degrees of freedom that must occur in any structured domain'. His illustrations include excerpts from Spearman and Thurstone on factor analysis, Chomsky on language, and Festinger, Schacter and Back on group structure.

This book is highly recommended for everyone who wishes to have a look at the field of mathematical psychology.

JOHN BROWN

Intermediate Statistics. By R. S. RODGER. Sydney: University Co-operative Bookshop. 1965. Pp. 168. 59*s.* 6*d.*

This textbook is written primarily for psychologists and social scientists. It is appropriately entitled 'intermediate', for it assumes from the outset that the reader already has some knowledge of the elementary concepts of descriptive statistics. It is a short book containing only 131 pages of actual text, but as it is concisely written it covers a fairly wide range of topics.

The first two chapters are introductory and deal with terminology and notions about hypothesis testing. The next two deal with simple analysis of variance models and are followed by a chapter on multiple contrasts in analysis of variance in which the methods proposed by Duncan, Tukey and Scheffé are discussed. In chapter 6 regression is briefly mentioned, in preparation for a discussion of trends in experimental data. Chapter 7 is entitled 'Frequency Counts' and the topics considered are the binomial theorem and test of association for contingency tables. The *pièce de résistance* of the book is chapter 8, entitled 'Multivariate Analysis'. It is the longest chapter, for the author feels that this is a topic about which psychologists will need to know a good deal in future. It begins with an introductory account of matrices, determinants and vector algebra. Matrix notation is then used to show how the univariate normal distribution generalizes into the multivariate normal. Next, multiple regression is presented in matrix terminology and several fairly well known multivariate tests of significance are given. These include the large sample test for the significance of a correlation matrix; a test for the difference between two variance–covariance matrices, and Hotelling's test for the difference between two vectors of means. The chapter ends with a brief account of principal component analysis, but reference to the more difficult topic of factor analysis is omitted.

A selection of exercises for students is added at the end of the text, and this is followed by 33 pages of statistical tables which include tables of the Studentized Range (needed for multiple comparison tests), in addition to the more usual ones.

The reviewer has mixed feelings about this book. There is no doubt that an 'intermediate' textbook on statistics for psychologists is needed, but the range of difficulty in this book varies too much. The chapters on analysis of variance tend to be sketchy and too elementary while the chapter on multivariate analysis is too condensed to serve as an introduction to the subject. The book should prove stimulating to teachers, but only the more industrious students are likely to benefit from it.

A. E. MAXWELL

Psychodiagnostic Testing: an Empirical Approach. By MOLLY HARROWER. Springfield, Illinois: C. C. Thomas. 1965. Pp. 90. $4.75.

In Part I, Professor Harrower asks four questions: 1. Is time in treatment related to success as assessed by the therapist? 2. Is age related to such 'success'? 3. Is sex related to such 'success'? 4. Is type of psychotherapy related to such 'success'? She answers 1 and 4 affirmatively and 2 and 3 negatively. Only the first of these conclusions is justifiable (leaving aside

some grave sampling problems) and this, of course, has very limited value. Even a psychotherapist might be forgiven for showing greater resistance to admitting failure after two years than two months.

We cannot be sure that classical analysts are better satisfied with their results *because* they are doing classical analysis, since the relation between outcome and length of treatment *within* type of therapy was not examined. The same applies to age and sex. Multivariate analysis seems necessary for studies of this nature.

In Part II, Prof. Harrower develops from test signs a homogeneous (low inter-test variability) and a heterogeneous (high variability) scale of personality endowment. In Part III, both scales are shown to be somewhat more significantly related to the therapists' assessment of outcome than is length of treatment: and multiple correlations are somewhat higher still, being 0·580 and 0·457 respectively. When, however, with a small subsample, each scale is related to outcome with therapist, method and time in treatment held constant (and the statistic is calculated), the association is not significant in either case. Furthermore, it would appear from charts I and II on pages 51 and 52 that there is an association, significant well beyond the 0·1% level of confidence, between good Personality Endowment and selection for Classical Analysis rather than Brief Therapy. The upshot seems likely to be that if you have Very Superior Personality Endowment and are without problems, your psychotherapist will probably elect to give you a Classical Analysis. This will probably last longer than would any of the rejected forms of treatment; but, at the end of this protracted treatment, your analyst will most likely be well pleased.

Despite this somewhat inglorious conclusion, Prof. Harrower is surely right to develop test scales and to relate them to psychotherapeutic outcome. It is much easier to scoff at the inadequacies of such studies than to carry out one that meets all scientific requirements. For the achievement of this ultimate goal it is probably better to have tried and partially lost than never to have tried at all. What we have gained is a greater awareness of the need to control, be it in psychotherapy or behaviour therapy studies, some such variable as Prof. Harrower calls 'mental health potential'.

In Part IV, this 'mental health potential' is related to success in the Unitarian ministry in a more convincing manner, since four independent judges were used to assess 'success'.

G. A. FOULDS

Educational Psychology. By DONALD R. GREEN. New Jersey: Prentice-Hall. 1964. Pp. vii + 120. 12*s*.

This is one of a series of books aiming to cover the field of psychology in short self-contained volumes each confined to a separate topic. Paperbacked series of this kind are now becoming fashionable on the other side of the Atlantic. They have the advantage of cheapness and flexibility. They are of convenient size and the treatment of a single topic is usually adequate for an introductory course. Furthermore, the student is able to select just those topics which are relevant to his particular studies.

In a book of this length the author has had to be highly selective and has adopted a very limited interpretation of 'educational psychology'. He has concentrated, therefore, on the highly important aspect of learning and teaching in the school environment. Theoretical discussions are firmly anchored to the real classroom situation. The text is relevant to the practice of education and is not open to the usual criticism of educators that books on the psychology of learning have little practical value. It is lucidly and concisely written, the material is well organized and a great deal of useful information is packed in a small space. Concepts discussed include intelligence, personality, psychological environment, meaning, set, attitude, motivation, instructional techniques and educational method. Chapters are devoted to the consideration not of student characteristics only, but also those of teachers. The linking theme adopted throughout the book is transfer of training. This concept is used in rather a broad manner, so that at one point the mere use of the techniques of one subject in another study is considered to be transfer of training. The book is written entirely from the American standpoint and includes some material, such as the effect of racial segregation of schools, which is quite irrelevant to the British educational system. The argument is well supported by research findings, again entirely American. There is a short, but useful, bibliography and index.

The balance of the text, however, might well be questioned. Only one page is devoted to exceptional children and but part of that to the slower learners. No other handicaps are mentioned, nor are the specific needs of the various age-groups. Creativity receives brief treatment, but the role of interest, as a motivational factor in learning, is almost overlooked.

The author complains of lack of psychological knowledge in the teachers and of teaching expertise in the psychologists. Even if we are to assume that this is general in the United States of America, which seems doubtful, we certainly cannot generalize in this way about the position in this country. His criticism that student teachers are taught a great deal in methods courses, but that very little of their time is given to the study of actual teaching techniques, is frequently echoed by head teachers here.

This is perhaps not one of the best volumes in this series. However, it can be stimulating and it may well prove useful in Colleges of Education and in postgraduate training courses, but amplification and discussion will be necessary if the student is not to be misled on some points. In courses for experienced teachers it could be the basis for some lively seminars. A. WALBRIDGE

Psychology and Teaching. By H. S. N. McFARLAND. 2nd ed. London: Harrap. 1965. Pp. 250. 12*s*. 6*d*.

This is the second edition of the book first published in 1958. The new edition has been revised by the inclusion of references in the text to recent educational reports and new legislation. In addition there is a new chapter on modern trends and problems.

The book is intended for the student receiving his initial teacher training and covers the usual range of topics such as learning and thinking, individual differences in ability and personality, development in childhood and adolescence, backwardness, delinquency and mental health. There are chapters on aesthetic and technical education and one each on social and biological views of man.

The emphasis of the book is not so much on a systematic exposition of researches and ideas in educational psychology as on the discussion of important practical issues in teaching—learning, assessment, adjustment, discipline—making reference to relevant psychological concepts and investigations. It is thus in the nature of an introductory book aiming at providing a basis of ideas and tempting students to read further in educational and child psychology. There is a useful bibliography at the end with selected books indicated which would prove 'a moderately easy next step' into various topics. The presentation is clear and readable and the young teacher is not likely to feel overtaxed with theory and technicalities.

However, since some students will not be tempted further, one could wish for a good deal more detail about some topics. For example, young teachers need more information and concepts to guide their observation and thinking about development in childhood and adolescence; they need more than is provided about the processes of learning and thinking and the growth of language. The new chapter on modern trends and problems contains brief accounts of important fields of inquiry which could usefully have been expanded. This chapter contains a short account of Piaget's work, a discussion of recent ideas about learning and intelligence, a section on programmed learning and one on social psychology and education. This is rather much for one short chapter and it would have been better if these ideas could have been expanded and included at relevant places in earlier chapters.

This book is scarcely sufficient as a basic text for courses in teacher training and would need a good deal of supplementing from other sources. It does, however, provide a readable introduction to a wide range of topics and would be useful reading for students in their initial training or as preliminary reading for teachers taking advanced courses in education who need a gentle introduction to fields they will have to explore more thoroughly.

RONALD GULLIFORD

Brit. J. Psychol. (1965), 56, 4, *pp.* 502–512
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review.)

ABRAHAMSEN, D. *The Psychology of Crime*. London: Tavistock. 1965. Pp. xii + 358. 15*s*.

ADINARYAN, S. P. *The Case for Colour*. London: Asia Publishing House. 1965. Pp. vii + 208. 25*s*.

AICHHORN, A. *Delinquency and Child Guidance*. New York: International Universities Press. 1965. Pp. 244. 45*s*.

ALEXANDER, C. *Notes on the Synthesis of Form*. Cambridge, Mass.: Harvard U.P.; London: O.U.P. 1964. Pp. 216. 54*s*.

ALIN, L. H. *Experimental Studies in Verbal Versus Figural Learning and Retention*. V. Ed. J. ELMGREN. Stockholm: Almquist and Wiksell. 1964. Pp. 216. 40 Sw. kr.

AMOS, J. R., BROWN, F. L. & MINK, O. G. *Statistical Concepts*. London: Harper and Row. 1965. Pp. ix + 125. 13*s*.

ANGYAL, A. *Neurosis and Treatment: Holistic Theory*. New York: Wiley. 1965. Pp. xxi + 328. 60*s*.

ANSBACKER, H. & ANSBACKER, R. R. *Superiority and Social Interest*. London: Routledge and Kegan Paul. 1965. Pp. xix + 432. 45*s*.

ARMSTRONG, C. N. & MARSHALL, A. J. *Intersexuality in Vertebrates including Man*. London: Academic Press. 1964. Pp. xi + 479. 90*s*.

ASRATIAN, E. *Compensatory Adaptations, Reflex Activity and the Brain*. Oxford: Pergamon. 1965. Pp. xv + 194. 80*s*.

ATKINSON, J. W. *An Introduction to Motivation*. University Series in Psychology. London: van Nostrand. 1964. Pp. xii + 335. 76*s*.

ATKINSON, R. C. (ed.). *Studies in Mathematical Psychology*. London: Oxford University Press. 1965. Pp. viii + 414. 92*s*.

AYER, A. J. *Man as a Subject for Science*. Auguste Comte Memorial Lecture No. 6. London: Athlone Press. 1964. Pp. 26. 5*s*.

BAEYER, W. R. V., HAFNER, H. & KISKER, K. P. *Psychiatrie der Verfolgten*. 1964. Berlin: Springer-Verlag. Pp. xii + 397. DM. 78.

BAKER, C. *Visual Capabilities in the Space Environment*. Oxford: Pergamon. 1965. Pp. vii + 203. 63*s*.

BANTON, M. *Roles*. London: Tavistock. 1965. Pp. ix + 224. 21*s*.

BARKER, R. G. & GUMP, P. V. *Big School, Small School. High School Size and Student Behavior*. Stanford: University Press; London: Oxford University Press. 1964. Pp. vii + 250. 48*s*.

BAROJA, JULIO C. *The World of the Witches*. Nature of Human Society Series. 1964. London Weidenfeld and Nicolson. Pp. xiv + 313. 36*s*.

BARRON, F. *et al. New Directions in Psychology*. Vol. II. New York: Holt, Rinehart and Winston. 1965. Pp. x + 422. $3.50.

BARTLEY, S. H. *Fatigue*. Springfield, Ill.: Thomas. 1965. Pp. xv + 96. $5.00.

BARTLETT, Sir FREDERICK. *Thinking: An Experimental and Social Study*. London: Unwin University Books. 1964. Pp. 203. 15*s*.

BASS, B. M. *Organizational Psychology*. Boston: Allyn and Bacon. 1965. Pp. xiii + 459. $8.95.

BENNISS, W. G., SCHEIN, E. H., BERLEW, D. E. & STEELE, F. I. *Interpersonal Dynamics. Essays on Human Interaction*. Homewood, Ill.: Dorsey Press. 1964. Pp. xv + 763. $8.50.

Berlyne, D. E. *Structure and Direction in Thinking*. New York and London: Wiley. 1965. Pp. xi + 378. 68s.

Bier, W. C. *Marriage. A Psychological and Moral Approach*. New York: Fordham University Press. 1965. Pp. viii + 280. $5.50

Bischof, L. J. *Interpreting Personality Theories*. New York and London: University Books. 1964. Pp. x + 304. $7.50.

Bloom, B. S. *Stability and Change in Human Characteristics*. London: Wiley. 1964. Pp. xiv + 238. 53s.

Bonner, H. *On Being Mindful of Man*. Boston: Houghton Mifflin. 1965. Pp. xlii + 254. 32s. 6d.

Bosselman, B. C., Rosenthal, I. M. & Schwarz, M. *Introduction to Developmental Psychiatry*. Springfield, Ill.: Thomas. Pp. 139. $6.50.

Brandon, O. *Christianity from Within*. London: Hodder and Stoughton. Pp. xii + 157. 18s.

Breer, P. E. & Locke, E. A. *Task Experience as a Source of Attitudes*. Homewood, Ill.: Dorsey Press. 1965. Pp. x + 280. $10.60.

Brenman, M. & Gill, M. M. *Hypnotherapy*. London: Wiley. 1964. Pp. 276. 13s.

Broadbent, D. E. *Behaviour*. University Paperbacks. London: Methuen. 1964. Pp. 215. 9s. 6d.

Brown, C. G. *Instrumentation with Semiconductors*. Springfield, Ill.: Thomas. 1964. Pp. xviii + 254. $10.50.

Brownfield, C. A. *Isolation*. New York: Random House. 1965. Pp. xii + 180. $1·95.

Buber, M. *The Knowledge of Man*. London: Allen and Unwin. 1965. Pp. 187. 28s.

Bullough, Vern L. *The History of Prostitution*. New York: University Books. 1964. Pp. x + 304. $7.50.

Burket, G. R. *A Study of Reduced Rank Models for Multiple Prediction*. Psychometric Monograph No. 12. Richmond, Virginia: William Byrd. 1964. Pp. xi + 66. $4.00.

Burrow, T. *Preconscious Foundations of Human Experience*. London: Basic Books. 1965. Pp. xxii + 164. 30s.

Burton, G. *Nurse and Patient*. London: Tavistock. 1965. Pp. xi + 220. 21s.

Caplan, G. *Principles of Preventive Psychiatry*. London: Tavistock. 1964. Pp. xi + 304. 42s.

Carroll, J. B. *Language and Thought*. Foundations of Modern Psychology Series. New Jersey: Prentice-Hall. 1964. Pp. x + 119. 12s. 6d.

Cassie, A., Fokkema, S. D. & Parry, J. B. *Aviation Psychology*. The Hague: Mouton. 1964. Pp. 125. fl. 17.50.

Cerda, E. *Una Psicologia de Hoy*. Barcelona: Herda. 1965. Pp. xvi + 709. No price given.

Chesser, E. *Unmarried Love*, London: Jarrolds. 1965. Pp. 160. 15s.

Chein, I., Gerard, D. L., Lee, R. S. & Rosenfeld, E. *Narcotics, Delinquency, and Social Policy*. London: Tavistock. 1964. Pp. x + 482. 48s.

Chombart de Lauwe, M.-J. & P.-H. *et al. La Femme dans la Société : Son Image dans Différents Milieux Sociaux*. Paris: Centre National de la Recherche Scientifique. 1963. Pp. 439. 32F.

Ch'u Chai, Winberg, C. (eds.) *I Ching. Book of Changes*. New York: University Books, Inc. 1965. Pp. ci + 418. $10.00.

CIBA Foundation Study Group No. 18. *Brain–Thyroid Relationships*. London: J. and A. Churchill. 1964. Pp. 117. 15s.

Coate, M. *Beyond all Reason*. London: Constable. 1964. Pp. x + 227. 21s.

Cohen, A. R. *Attitude Change and Social Influence*. London: Basic Books. 1964. Pp. xv + 156. 18s.

Cohen, J. *Behaviour in Uncertainty and its Social Implications*. London: Allen and Unwin. 1964. Pp. 207. 30s.

Cohen, S. *Drugs of Hallucination*. London: Secker and Warburg. 1965. Pp. xviii + 268. 30s.

Collins, B. L. & Guetzkow, H. *A Social Psychology of Group Processes for Decision Making.* London: Wiley. 1964. Pp. x + 254. 47*s.*

Colodny, R. G. (ed.). *Frontiers of Science and Philosophy.* London: Allen and Unwin. 1964. Pp. 288. 28*s.*

Coombs, C. H. *A Theory of Data.* New York: Wiley. 1964. Pp. xviii + 585. 113*s.*

Crane, E. *Marketing Communications.* London: Wiley. 1965. Pp. xiv + 569. 68*s.*

Cronbach, L. J. & Gleser, G. C. *Psychological Tests and Personnel Decisions.* Urbana, Ill.: University of Illinois Press. 1965. Pp. viii + 347. $7.95.

Das, S. S. *General Psychology.* London: Asia Publishing House. 1964. Pp. xi + 220. 30*s.*

David, H. P. *International Resources in Clinical Psychology.* Maidenhead: McGraw-Hill. 1964. Pp. xiii + 236. 58*s.*

Davis, J. *et al. Tables for Yules Q Association Coefficient for Pairs of Percentages.* Chicago: Norc. 1965. Pp. 24. $2.00.

Dember, W. *Visual Perception : The Nineteenth Century.* New York: University Books. 1965. Pp. xii + 22. Cloth, 38*s.*, Paper, 23*s.*

Dennis, W. *The Hopi Child.* London: Wiley. 1965. Pp. vii + 200. 12*s.*

Dinkmeyer, D. C. *Child Development.* London: Prentice-Hall. 1965. Pp. xi + 434. 56*s.*

Durkin, H. E. *The Group in Depth.* New York: International Universities Press. 1965. Pp. xii + 378. 76*s.*

Easton, D. *A Systems Analysis of Political Life.* London: Wiley. 1965. Pp. xvi + 507. 68*s.*

Edwards, A. L. *Expected Values of Discrete Random Variables and Elementary Statistics.* London: Wiley. 1964. Pp. x + 146. Cloth 45*s.*, Paper 30*s.*

Eiferman, R. *Studies in Psychology.* Jerusalem: Magnes Press. 1965. Pp. 230. $6.00.

Eissler, K. R. *Medical Orthodoxy and the Future of Psychoanalysis.* New York: International Universities Press. 1965. Pp. x + 592. 108*s.*

Evans, R. I. *Conversations with Carl Jung, and Reactions from Ernest Jones.* London: Van Nostrand. 1964. Pp. viii + 173. 14*s.*

Eysenck, H. J. *Smoking, Health and Personality.* London: Weidenfeld and Nicolson. 1965. Pp. 166. 25*s.*

Eysenck, H. J. & Rachman, S. *The Causes and Cures of Neurosis.* London: Routledge and Kegan Paul. 1965. Pp. xii + 318. 28*s.*

Faure, H. *Hallucinations et Réalité Perceptive.* Paris: Presses Universitaires de France. 1965. Pp. 265. F. 18.

Festinger, L. *Conflict, Decision and Dissonance.* London: Tavistock. 1964. Pp. vii + 163. 37*s.* 6*d.*

Ferguson, G. A. *Nonparametric Trend Analysis.* Montreal: McGill University Press. 1965. Pp. 61. $1.95.

Fleishman, E. A. *The Examiner's Manual for the Basic Fitness Tests.* New Jersey: Prentice-Hall. 1964. Pp. iii + 60. No price given.

Fleishman, E. A. *The Structure and Measurement of Physical Fitness.* New Jersey: Prentice-Hall. 1964. Pp. xii + 205. No price given.

Fleming, C. M. *Teaching : A Psychological Analysis.* University Paperbacks. London: Methuen. 1964. Pp. xiii + 291. 12*s.* 6*d.*

Foulds, G. A. *Personality and Personal Illness.* London: Tavistock. 1965. Pp. xi + 344. 55*s.*

Foulkes, S. H. *Therapeutic Group Analysis.* London: Allen and Unwin. 1964. Pp. 320. 42*s.*

Fraisse, P. *The Psychology of Time.* London: Eyre and Spottiswoode. 1964. Pp. 343. 30*s.*

Freeman, L. C. *Elementary Applied Statistics.* London: Wiley. 1965. Pp. viii + 298. 53*s.*

FRIED, E. *Artistic Productivity and Mental Health.* Springfield, Ill.: Thomas. 1964. Pp. xiv + 177. $6.50.

FRIEDLANDER, E. *Psychology in Scientific Thinking.* New York: Philosophical Library. 1965. Pp. 142. $4.75.

FULTON, R. *Death and Identity.* New York: Wiley. 1965. Pp. xv + 415. 53*s*.

GALLAGHER, J. R. & HARRIS, H. I. *Emotional Problems of Adolescents.* London: Oxford University Press. 1964. (2nd edition). Pp. x + 210. 35*s*.

GEIST, H. *A Child Goes to Hospital.* Springfield, Ill.: Thomas. 1965. Pp. ix + 112. $5.75.

GELLHORNE, E. & LOOFBOURROW, G. N. *Emotions and Emotional Disorders.* New York: Harper and Row. 1965. Pp. xii + 496. $12.00.

GLUECK, S. & GLUECK, E. *Ventures in Criminology: Selected Recent Papers.* London: Tavistock. 1964. Pp. xiv + 371. 65*s*.

GLUCKMAN, M. (ed.) *Closed Systems and Open Minds. The Limits of Naïvety in Social Anthropology.* Edinburgh: Oliver and Boyd. 1964. Pp. viii + 274. 42*s*.

GOERTZEL, V. & GOERTZEL, M. *Cradles of Eminence.* London: Constable. 1965. Pp. xv + 362. 42*s*.

GORER, G. *Death, Grief and Mourning in Contemporary Britain.* London: Cresset Press. 1965. Pp. viii + 184. 35*s*.

GOULD, J. & KOLB, W. L. (eds.) *A Dictionary of the Social Sciences.* London: Tavistock. 1964. Pp. xvi + 761. 6½ gns.

GOULD, R. T. *Oddities: A Book of Unexplained Facts.* New York: University Books. 1965. Pp. xi + 220. $5.00.

GOULD, R. T. *Enigmas: Another Book of Unexplained Facts.* New York: University Books. 1965. Pp. xi + 240. $5.00.

GOUREVITCH, V. *Statistical Methods: A Problem-solving Approach.* Boston: Allyn and Bacon. 1965. Pp. ix + 310. $7.75.

GREEN, M. R. (ed.). *Interpersonal Psychoanalysis. The Selected Papers of Clara M. Thompson* London: Basic Books. 1965. Pp. xii + 398. 50*s*.

GROOT, A. D. D. *St Nicholas.* The Hague: Mouton. 1965. Pp. 211. $5.90.

GROOT, A. D. D. *Thought and Choice in Chess.* The Hague: Mouton. 1965. Pp. xvi + 463. $10.00.

GROSSMAN, S. *The Wild Analyst: The Life and Work of Groddeck.* London: Barrie and Rockliffe. 1965. Pp. 211. 25*s*.

GRYGIER, T. *et al. Criminology in Transition. Essays in Honour of Mannheim.* London: Tavistock. 1965. Pp. xix + 308. 45*s*.

GUILFORD, J. P. *Fundamental Statistics in Psychology and Education.* Maidenhead: McGraw-Hill. 1965. Pp. 606. 68*s*.

GUION, R. M. *Personnel Testing.* Maidenhead: McGraw-Hill. 1965. Pp. xi + 585. 76*s*.

GUSTAFSON, S. R. *The Pediatric Patient.* London: Pitman. 1964. Pp. 160. 35*s*.

HADFIELD, J. A. *Psychology and Morals. An Analysis of Character.* London: Methuen. 1964. (16th edn.). Pp. ix + 246. 9*s*. 6*d*.

HALLORAN, J. D. *The Effects of Mass Communication.* Leicester: University Press. 1964. Pp. 83. 7*s*. 6*d*.

HARKER, J. E. *The Physiology of Diurnal Rhythms.* Cambridge Monographs in Experimental Biology. London: Cambridge University Press. 1964. Pp. 114. 20*s*.

HARPER, R. J. C., ANDERSON, C. C., CHRISTENSEN, C. M. & HUNKA, S. M. *The Cognitive Processes: Readings.* New Jersey: Prentice-Hall. 1964. Pp. xii + 717. 72*s*.

HARRIS, E. E. *The Foundations of Metaphysics in Science.* London: Allen and Unwin. 1965. Pp. 512. 63*s*.

Hartman, H. *Essays on Ego Psychology. Selected Problems in Psychoanalytic Theory.* London: Hogarth Press. 1964. Pp. xv + 492. 75*s.*

Haworth, M. R. *Child Psychotherapy.* London: Basic Books. 1964. Pp. xv + 459. 50*s.*

Heckman, B. & Fried, R. *A Manual of Laboratory Studies in Psychology.* London: Oxford University Press. 1965. Pp. xiii + 207. 20*s.*

Helson, H. *Adaptation-Level Theory. An Experimental and Systematic Approach to Behaviour.* New York and London: Harper and Row. 1964. Pp. xvii + 732. 81*s.*

Hendrick, I. *Psychiatry Education Today.* New York: International Universities Press. 1965. Pp. vii + 110. 27*s.*

Hersey, R. *The Myth of the Inscrutable Oriental.* Ocean City: Littoral Development. 1965. Pp. vii + 61. $1.00.

Hess, W. R. *The Biology of Mind.* London: University of Chicago Press. 1964. Pp. xii + 203. 45*s.*

Higgin, G. & Jessop, N. *Communications in the Building Industry.* London: Tavistock. 1965. Pp. 125. 12*s.* 6*d.*

Hill, Winfred F. *Learning: A Survey of Psychological Interpretations.* London: Methuen. University Paperbacks. 1964. Pp. x + 227. 9*s.* 6*d.*

Himwich, W. & Schade, J. (eds.). *Horizons in Neuro-psycho-pharmacology.* London: Elsevier. 1965. Pp. 357. 100*s.*

Hochberg, J. E. *Perception.* Foundations of Modern Psychology Series. New Jersey: Prentice-Hall. 1964. Pp. x + 118. 18*s.*

Hollander, E. P. *Leaders, Groups and Influence.* London: Oxford University Press. 1965. Pp. xi + 256. 40*s.*

Holt, A. G. *Handwriting in Psychological Interpretation.* Springfield, Ill.: Thomas. 1965. Pp. xiii + 262. $10.50.

Homer, W. I. *Seurat and the Science of Painting.* Cambridge, Mass.: M.I.T. Press. 1964. Pp. xvi + 327. 94*s.*

Horrobin, D. *Communication Systems of the Body.* London: Basic Books. 1965. Pp. viii + 214. 21*s.*

Hovland, C. I. *et al. Experiments on Mass Communication.* London: Wiley. 1965. Pp. x + 345. 15*s.*

Howe, E. G. *Cure or Heal.* London: Allen and Unwin. 1965. Pp. 236. 28*s.*

Hultzen, L. S., Allen, J. H. D. & Miron, M. S. *Tables of Transitional Frequencies of English Phonemes.* Illinois: University of Illinois Press. 1964. Pp. 49. $5.00.

Hutt, M. L. & Gibby, R. G. *The Mentally Retarded Child.* Boston: Allyn and Bacon. 1965. Pp. xiv + 441. $7.95.

Hyman, R. *The Nature of Psychological Inquiry.* Foundations of Modern Psychology Series. New Jersey: Prentice-Hall. 1964. Pp. xi + 116. 12*s.* 6*d.*

Isaacson, R. L., Hutt, L. M. & Blum, M. L. *Psychology: the Science of Behaviour.* New York: Harper and Row. 1965. Pp. ix + 374. 52*s.*

Itoh, H. *The Essence of the Universe and the Mystery of the Spirit.* Tokyo: Gakkai. 1965. Pp. vi + 37. No price given.

Jackson, D. D. *Myths of Madness.* New York: Macmillan. 1964. Pp. 178. 35*s.*

Jacobsen, E. *Anxiety and Tension Control. A Physiologic Approach.* Philadelphia: J. B. Lippincott. London: Pitman. 1964. Pp. xvi + 216. 32*s.*

Jacobson, E. *The Self and the Object World.* London: Hogarth Press. 1965. Pp. xii + 250. 35*s.*

Jaspers, K. *The Nature of Psychotherapy.* Manchester University Press. 1964. Pp. x + 49. 7*s.* 6*d.*

Johnson, R. C. & Medinnus, G. *Child Psychology: Behaviour and Development.* New York: Wiley. 1965. Pp. x + 556. 53*s.*

Jung, C. *Answer to Job.* London: Hodder and Stoughton. 1965. Pp. xviii + 192. 6*s.*

Kadis, A. L. & Winick, C. (eds.). *Topical Problems of Psychotherapy. Group Psychotherapy Today.* Vol. 5. Basel: Karger. 1965. Pp. vi + 256. Fr/DM 39.

Kahn, S. *Psychodrama Explained.* New York: Philosophical Library. 1964. Pp. vi + 77. $3.00.

Kantor, M. B. *Mobility and Mental Health.* Springfield, Ill.: Thomas. 1965. Pp. xix + 247. $10.50.

Kastenbaum, R. *New Thoughts on Old Age.* New York: Springer. 1964. Pp. xii + 333. $7.50.

Kellmer Pringle, M. L. *Deprivation and Education.* London: Longmans. 1965. Pp. 307. 42*s*. 6*d*.

Kelly, H. P. Memory Abilities: A Factor Analysis. *Psychometric Monographs*, No. 11. Richmond: William Byrd. 1964. Pp. xi + 53. $3.00.

Kendler, H. *Basic Psychology.* London: Methuen. 1965. Pp. xii + 750. 60*s*.

Kiell, N. *Psychiatry and Psychology in the Visual Arts.* Wisconsin: University Press. 1965. Pp. xiv + 250. $6.50.

King, D. *Training within the Organization.* London: Tavistock. 1964. Pp. xxii + 274. 38*s*.

Kisker, G. W. *The Disorganized Personality.* Maidenhead: McGraw-Hill. 1964. Pp. 631. 66*s*.

Kissen, D. M. & LeShan, L. L. (eds.). *Psychosomatic Aspects of Neoplastic Disease.* London: Pitman. 1964. Pp. 231. 21*s*.

Kissle, P. & Barrucand, D. *Placébos et Effet Placébo en Médecine.* Paris: Masson. 1964. Pp. 240. 38F.

Kleinsorge, H. & Klunbies, G. *Techniques of Relaxation.* Bristol: Wright. 1965. Pp. 68. 35*s*.

Klineberg, O. *The Human Dimension in International Relations.* London: Holt, Rinehart and Winston. 1964. Pp. viii + 173. 22*s*.

Kogan, N. & Wallach, M. A. *Risk Taking. A Study in Cognition and Personality.* London: Holt, Rinehart and Winston. 1964. Pp. 278. $8.00.

Kornhauser, A. *Mental Health of the Industrial Worker.* London: Wiley. 1965. Pp. xi + 354. 60*s*.

Kornrich, M. *Psychological Test Modifications.* Springfield, Ill.: Thomas: 1965. Pp. xii + 265. $8.75.

Landis, C. *Varieties of Psychopathological Experience.* (Ed. by Mettler, F.). New York: Holt, Rinehart and Winston. 1964. Pp. xviii + 484. $8.75.

Lang, A. *Über Zwei Teilsysteme der Persönlichkeit.* Berne: Hans Huber. 1964. Pp. 147. Fr/DM 18.

Lawson, J. C. *Modern Greek Folklore and Ancient Greek Religion.* New York: University Books. 1964. Pp. xxii + 620. $10.00.

Leary, T., Metzner, R. & Alpert, R. *The Psychedelic Experience. A Manual Based on the Tibetan Book of the Dead.* New York: University Publications. 1964. Pp. 159. $5.00.

Lerbinger, O. & Sullivan, A. J. (eds.). *Information, Influence and Communication.* London: Basic Books. 1965. Pp. xii + 513. 60*s*.

Levine, M. & Spivack, G. *The Rorschach Index of Repressive Style.* Springfield, Ill.: Thomas. 1964. Pp. xvi + 164. $7.50.

Levinson, D. J. & Gallagher, E. B. *Patienthood in the Mental Hospital.* Boston: Houghton Mifflin. 1964. Pp. xvii + 265. $5.95.

Lewin, L. *Phantastics, Narcotic and Stimulating Drugs.* London: Routledge and Kegan Paul. 1964. Pp. xvii + 335. 30*s*.

Lindzey, G. & Hall, C. S. *Theories of Personality.* London: Wiley. 1965. Pp. xii + 543. 53*s*.

Ljung, B. O. *The Adolescent Spurt in Mental Growth.* Stockholm: Almquist and Wiskell. 1965. Pp. xxx + 350. Sw. Kr. 48.

London, P. *The Modes and Morals of Psychotherapy.* New York: Holt, Rinehard and Winston. 1964. Pp. ix + 278. $5.50.

Lowenthal, Marjorie F. *Lives in Distress. The Paths of the Elderly to the Psychiatric Ward.* London: Basic Books. 1964. Pp. xx + 266. 30*s*.

Luce, R. D., Bush, R. R. & Galanter, E. *Readings in Mathematical Psychology.* Vol. II. London: Wiley. 1965. Pp. ix + 568. 68*s*.

Luce, R. D., Bush, R. R. & Galanter, E. *Handbook of Mathematical Psychology.* Vol. III. London: Wiley. 1965. Pp. ix + 537. 90*s*.

Lyons, J. *A Primer of Experimental Psychology.* London: Harper and Row. 1965. Pp. xii + 322. 26*s*.

McCurdy, H. G. *Personality and Science. A Search for Self-Awareness.* New Jersey: van Nostrand. 1965. Pp. viii + 151. $1.45.

McGrath, J. E. *Social Psychology. A Brief Introduction.* New York: Holt, Rinehart and Winston. 1964. Pp. ix + 166. 22*s*.

Macnab, F. A. *Estrangement and Relationship. Experience with Schizophrenia.* London: Tavistock. 1965. Pp. xix + 299. 45*s*.

McNeil, E. R. *et al. The Nature of Human Conflict.* New Jersey: Prentice-Hall. 1965. Pp. xvi + 315. No price given.

Maier, N. R. F. *Psychology in Industry.* London: Harrap. 1965. Pp. xiv + 718. (3rd edn.). 50*s*.

Maier, N. R. F. & Schneirla, T. C. *Principles of Animal Psychology.* London: Constable. 1965. Pp. xvi + 683. 24*s*.

Majumdar, S. K. *Introduction to Yoga: Principles and Practices.* New York: University Books. 1964. Pp. 318. $7.50.

Mandler, G. & Kessen, W. *The Language of Psychology.* (Science edition.) London: Wiley. 1964. Pp. xviii + 301. 15*s*.

Marienfeld, H. *Simulationstechnik. Allgemeine Grundlagen und Anwendung in der Starrfluglertechnik.* Düsseldorf: Vdi-Verlag. 1964. Pp. viii + 888. No price given.

Marmor, J. *Sexual Inversion.* London: Basic Books. 1965. Pp. vii + 358. 60*s*.

Martin-Santos. *Libertad, Temporalidad y Transferencia en el Psicoanalisis Existencial.* Editorial Seis Barral. S. A. 1964. Pp. xv + 204. No price given.

Mashhour, M. *Psychophysical Relations in the Perception of Velocity.* Stockholm: Almquist and Wiksell. 1964. Pp. 176. Sw.kr. 22.

Mead, M. *Anthropology: A Human Science.* Insight 22. London: van Nostrand. 1964. Pp. viii + 260. 15*s*. 6*d*.

Mead, M. *Continuities in Cultural Evolution.* London: Yale University Press. 1964. Pp. xxii + 471. 63*s*.

Meade, J. E. & Parkes (eds.). *Biological Aspects of Social Problems.* Edinburgh: Oliver and Boyd. 1965. Pp. x + 226. 50*s*.

Medical Research Council. *Abnormalities of the Sex Chromosome Complement in Man.* Special Report Series, no. 305. London: H.M.S.O. 1964. Pp. viii + 239. 27*s*. 6*d*.

Meerloo, J. A. M. & Nelson, M. C. *Transference and Trial Adaptation.* Springfield, Ill.: Thomas. 1965. Pp. vii + 155. $6.50.

Meggitt, M. J. *The Lineage System of the Mac–Enga of New Guinea.* Edinburgh: Oliver and Boyd. 1965. Pp. xviii + 297. 70*s*.

Meister, D. & Rabidean, G. *Human Factors Evaluation in System Development.* London: Wiley. 1965. Pp. xii + 307. 75*s*.

Mendelssohn, F. V. *This is Psychiatry.* New York: Frankin Watts. 1964. Pp. xii + 237. $7.95.

Messick, S. & Brayfield, A. H. *Decision and Choice: Contributions of Sidney Siegal.* Maidenhead: McGraw-Hill. 1964. Pp. xi + 298. 69*s*. 6*d*.

Miller, D. *Growth to Freedom. The Psychosocial Treatment of Delinquent Youth.* London: Tavistock. 1964. Pp. xxv + 224. 30*s*.

Miner, J. B. *Studies in Management Education.* New York: Springer. 1965. Pp. viii + 239. Pp. 63. $7.50.

Money, J. (ed.). *Sex Research: New Developments.* New York: Holt, Rinehart and Winston. 1965. Pp. xii + 260. $3.50.

Morrisett, L. W. & Vinsohaler, J. Mathematical learning. *Monograph of Society for Research in Child Development*, **30**, no. 1. 1965. Pp. 150. $6.00.

Morton, G. M. *The Arts of Costume and Personal Appearance* (3rd ed.). New York and London: Wiley. 1964. Pp. 319. 60*s*.

Moustaka, C. *The Internal Migrant.* Athens: Social Sciences Centre. 1965. Pp. xii + 105. No price given.

Mueller, C. G. *Sensory Psychology.* London: Prentice-Hall. 1965. Pp. viii + 120. Cloth 32*s*., paper 12*s*. 6*d*.

Muensterberger, W. & Axelrad, S. (eds.). *The Psychoanalytic Study of Society.* (Vol. iii). New York: International Universities Press. 1965. Pp. 408. 68*s*.

Mumford, E. *Memoirs of a Big Bunny.* New York: Exposition Press. 1964. Pp. 159. $3.50.

Munn, N. L. *The Evolution and Growth of Human Behavior.* Boston: Houghton Mifflin. 1965. (2nd edn.) Pp. xi + 594. $8.50.

Murray, E. J. *Motivation and Emotion.* Foundations of Psychology Series. New Jersey: Prentice-Hall. 1964. Pp. ix + 118. 12*s*. 6*d*.

Nash, E. M., Jessner, L. & Abse, D. W. (eds.). *Marriage Counselling in Medical Practice. A Symposium.* London: Oxford University Press. 1965. Pp. xvi + 368. 64*s*.

Neubauer, P. B. *Early Childhood Education.* Springfield, Ill.: Thomas. 1965. Pp. viii + 149. $6.50.

Neustadt, I. *Teaching Sociology: An Inaugural Lecture.* Leicester: University Press. 1965. Pp. 23. 3*s*.

Nunberg, H. *Practice and Theory of Psychoanalysis.* Vol. ii. New York: International Universities Press. 1965. Pp. iii + 219. 40*s*.

Osborn, R. *Marxism and Psycho-Analysis.* London: Barrie and Rockliffe. 1965. Pp. xix + 160. 21*s*.

Palermo, D. S. & Jenkins, J. J. *Word Association Norms. Grade School Through College.* London: Oxford University Press. Minnesota University Press. 1964. Pp. x + 469. 60*s*.

Pearson, L. *The Use of Written Communications in Psychotherapy.* Springfield, Ill.: Thomas. 1965. Pp. xiii + 65. $4.50.

Peck, P. F. & Havighurst, R. J. *The Psychology of Character Development.* New York: Wiley. 1965. Pp. xix + 267. 12*s*.

Peiper, A. *Cerebral Function in Infancy and Childhood.* (3rd edn.) (Trs. Nagler, B. & H.) New York: Consultants Bureau. 1964. Pp. xii + 683. $25.00.

Phillips, H. P. *Thai Peasant Personality.* London: Cambridge University Press. 1964. Pp. xii + 231. 48*s*.

Phillips, M. *Small Social Groups in England.* London: Methuen. 1965. Pp. 318. 15*s*.

Post, F. *The Clinical Psychiatry of Late Life.* Oxford: Pergamon. 1965. Pp. ix + 173. 21*s*.

Presthus, R. *Behavioural Approaches to Public Administration.* Alabama: University of Alabama. 1965. Pp. 158. $3.95.

Priestley, J. B. *Man and Time.* London: Allen. 1964. Pp. 319. 70*s*.

Quay, H. C. *Juvenile Delinquency.* London: van Nostrand. 1965. Pp. ix + 350. 55*s*.

Raphael, A. *Goethe and the Philosopher's Stone.* London: Routledge and Kegan Paul. 1965. Pp. xiv + 273. 35*s*.

Ratliff, F. *Mach Bands.* San Francisco: Holden Day. 1965. Pp. xiii + 365. $15.25.

Reed, C. F., Alexander, I. E. & Tomkins, S. S. (eds.). *Psychopathology: A Source Book.* New York: Wiley. 1965. Pp. xii + 803. 23*s*.

Reed, E. W. & Reed, S. C. *Mental Retardation: A Family Study.* London: Saunders. 1965. Pp. v + 719. 115*s*. 6*d*.

Reger, R. *School Psychology.* Springfield, Ill.: Thomas. 1965. Pp. xv + 213. $7.75.

RICE, A. K. *Learning for Leadership*. London: Tavistock. 1965. Pp. xii + 200. 30*s*.

RICHARDSON, S. A., DOHRENWEND, B. S. & KLEIN, D. *Interviewing: Its Forms and Functions*. London: Basic Books. 1965. Pp. vii + 380. 40*s*.

RIEFF, P. *Freud—The Mind of the Moralist*. London: Methuen. 1965. Pp. xxiii + 397. 18*s*.

RIMLAND, B. *Infantile Autism*. London: Methuen. 1965. Pp. xi + 282. 36*s*.

ROACH, J. H. L., SHERMAN, B. R. & ROACH, P. P. *Experiments in General Psychology*. New York: Harper and Row. 1965. Pp. ix + 192. 28*s*.

ROBINSON, H. B. & ROBINSON, N. M. *The Mentally Retarded Child: A Psychological Approach*. New York: McGraw-Hill. 1965. Pp. xvi + 639. 84*s*.

ROSEN, E. & GREGORY, I. *Abnormal Psychology*. London: Saunders. 1965. Pp. vi + 553. 58*s*.

ROSENFELD, H. A. *Psychotic States. A Psychoanalytical Approach*. London: Hogarth Press. 1965. Pp. 263. 42*s*.

ROSS, STRANGE. *Logical Foundations of Psychological Measurement*. Denmark: Munksgaard. 1964. Pp. 152. 36 Danish Kroner.

ROUCEK, J. S. *The Difficult Child*. New York: Philosophical Library. 1964. Pp. viii + 292. $7.50.

ROYCE, J. R. *Psychology and the Symbol*. New York: Random House. 1965. Pp. 117. $1.65.

SAMPSON, E. E. *Approaches, Contests and Problems of Social Psychology: A Book of Readings*. New Jersey: Prentice-Hall. 1964. Pp. ix + 576. 72*s*.

SANFORD, F. H. & CAPALDI, E. J. *Advancing Psychological Science*, Vols. I, II and III. London: Prentice-Hall. 1964. 16*s*.

SANFORD, N. *College and Character*. London: Wiley. 1964. Pp. xv + 308. 38*s*.

SARGENT, S. S. & STAFFORD, K. R. *Basic Teaching of the Great Psychologists*. New York: Doubleday. 1965. Pp. xv + 382. $1.45.

SARTRE, J.-P. *The Problem of Method* (tr. Barnes, H.). London: Methuen. 1964. Pp. xxxv + 181. 30*s*.

SCHNEIDER, A. A. *Personality Dynamics and Mental Health. Principle of Adjustment and Mental Hygiene*. New York: Holt, Rinehart and Winston. 1965. Pp. xii + 546. No price given.

SCHNEIDER, H. *Über den Autismus*. Monographien aus dem Gesamtgebiete der Neurologie und Psychiatrie. Berlin: Springer Verlag. 1964. Pp. 47. DM. 14.80.

SCHEIN, E. H. *Organisational Psychology*. New Jersey: Prentice-Hall. 1965. 12*s*. 6*d*.

SCHEIN, E. H. & BENNIS, W. G. *Personal and Organisational Change through Group Methods: the Laboratory Approach*. London: Wiley. 1965. Pp. x + 376. 62*s*.

SCHOLEN, G. G. *On the Kabbalah and its Symbolism*. London: Routledge and Kegan Paul. 1965. Pp. v + 216. 30*s*.

SCHULTZ, D. P. *Sensory Restriction: Effects on Behaviour*. New York: Academic Press. 1965. Pp. viii + 216. $7.50.

SEARLES, H. F. *Collected Papers on Schizophrenia and Related Subjects*. London: Hogarth Press. 1965. Pp. 797. 105*s*.

SHAFER, J. N. *Laboratory Exercises in Psychology*. New York: Holt, Rinehart and Winston. 1965. Pp. v + 204. $2.95.

SHAW, O. L. *Maladjusted Boys*. London: Allen and Unwin. 1965. Pp. 168. 28*s*.

SHERIF, C. W., SHERIF, M. & NEBERGALL, R. E. *Attitude and Attitude Change*. London: Saunders. 1965. Pp. xxi + 264. 58*s*.

SHERIF, M. & SHERIF, C. W. *Reference Groups. Exploration into Conformity and Deviation of Adolescents*. New York and London: Harper and Row:. 1964. Pp. xiv + 370. 45*s*.

SIEGAL, S. *Choice, Strategy and Utility*. Maidenhead: McGraw-Hill. 1964. Pp. x + 178. 46*s*.

SLAMA-GAZACU, T. *Communicarea in Procesul Muncii*. Bucuresti: Editura Stinntifica. 1964. Pp. 301. Lei 12.

SMITH, D. D. *Mammalian Learning and Behaviour*. London: Saunders. 1965. Pp. ix + 174. 45*s*. 6*d*.

SMITH, M. B., BRUNER, J. S. & WHITE, R. W. *Opinions and Personality*. London: Wiley. 1964. Pp. 284. 19*s*.

STAATS, A. W. *Human Learning. Studies Extending Conditioning Principles to Complex Behaviour*. New York: Holt, Rinehart and Winston. 1964. Pp. xv + 520. $8.75.

STERN, P. J. *The Abnormal Person and His World*. London: van Nostrand. University Series in Psychology. 1964. Pp. xi + 239. 40*s*.

STEVENS, H. A. & HEBER, R. (eds.). *Mental Retardation. A Review of Research*. Chicago and London: University of Chicago Press. 1965. Pp. xii + 502. 90*s*.

STOKES, A. *The Invitation in Art*. London: Tavistock. 1965. Pp. xxxiii + 67. 18*s*.

STRANGE, J. R. *Abnormal Psychology*. Maidenhead: McGraw-Hill. 1965. Pp. vi + 488. 64*s*.

STRASSER, S. *Phenomenologie und Erfahrungswissenschaft vom Menschen*. Berlin: de Gruyter. 1964. Pp. xv + 313. DM. 38.

STUDIES OF THE WARTEGG DRAWING COMPLETION TEST. Takala, M. Helsinki: Soumalainen Tiedatemia. Pp. 102. 1964. No price given.

SWETS. J. S. *Signal Detection and Recognition by Human Observers*. London: Wiley. 1964. Pp. xi + 702. 132*s*.

TAJFEL, H. & DAWSON, J. L. *Disappointed Guests*. London: Oxford University Press. 1965. Pp. 158. Paper 15*s*., cloth 25*s*.

TALLAND, G. A. *Deranged Memory*. New York: Academic Press. 1965. Pp. xi + 356. 76*s*.

TAYLOR, C. *The Explanation of Behaviour*. London: Routledge and Kegan Paul. 1964. Pp. ix + 278. 40*s*.

TAYLOR, C. W. *Widening Horizons in Creativity*. London: Wiley. 1965. Pp. xix + 466. 68*s*.

TEEVAN, R. C. & BIRNEY, R. C. *Theories of Motivation in Learning*. Insight 19. London: van Nostrand. 1964. Pp. vi + 215. 14*s*.

TEEVAN, R. C. & BIRNEY, R. C. *Theories of Motivation in Personality and Social Psychology*. Insight 20. London: van Nostrand. 1694. Pp. vi + 186. 14*s*.

TOCH, T. *The Social Psychology of Social Movements*. Indianapolis: Bobs-Merrill. 1965. Pp. viii + 257. $6.50.

TOMKINS, S. S. & IZARD, C. E. (eds.). *Affect, Cognition and Personality. Empirical studies*. New York: Springer. 1965. Pp. xii + 464. $10.00.

TYLER, L. E. *The Psychology of Human Differences*. New York: Appleton Century Crofts. 1965. (3rd edn.) Pp. viii + 572. $6.75.

VESEY, G. N. A. *The Embodied Mind*. London: Allen and Unwin. 1965. Pp. 114. 20*s*.

VICKERS, J. *The Art of Judgement. A Study of Policy Making*. London: Chapman and Hall. 1965. Pp. 242. 25*s*.

VROOM, V. *Work and Motivation*. London: Wiley. 1964. Pp. ix + 331. 52*s*.

WAPNER, S. *et al. The Body Percept*. New York: Random House. 1965. Pp. x + 112. $1.65.

WARREN, J. M. *et al.* (eds.). *The Frontal Granular Cortex and Behaviour*. New York and London: McGraw-Hill. 1964. Pp. x + 492. 140*s*.

WATSON, R. I. *Psychology of the Child*. London: Wiley. 1965. (2nd edn). Pp. ix + 635. 57*s*.

WEBER, M. *The Sociology of Religion*. London: Methuen. 1965. Pp. lxvii + 308. 30*s*.

WECKROTH, J. *Studies in Colour After-Images*. Helsinki: Academia Scientiarum Fennica. 1964. Pp. 110. No price given.

WEISSMANN, P. *Creativity in the Theatre*. London: Basic Books. 1965. Pp. x + 275. 30*s*.

WEITZ, H. *Behaviour Change Through Guidance*. London: Wiley. 1964. Pp. xii + 250. 49*s*.

WHITAKER, D. S. & LIEBERMAN, M. A. *Psychotherapy Through the Group Process*. London: Prentice-Hall. 1964. Pp. x + 305. 68*s*.

WHITTAKER, J. O. *Introduction to Psychology* and *Students Manual*. London: Saunders. 1965. Pp. xxiii + 631. 52*s*. 6*d*.

WILLETT, T. C. *Criminal on the Road. A Study of Serious Motoring Offences and Those Who Commit Them*. London: Tavistock. 1964. Pp. xv + 343. 42*s*.

WILLIAMS, M. *Mental Testing in Clinical Practice*. Oxford: Pergamon. 1965. Pp. xix + 177. 21*s*.

WILSON, I. G. & WILSON, M. E. *Information, Computers and Systems Design*. London: Wiley. 1965. Pp. xx + 341. 95*s*.

WILSON, R. N. *The Arts in Society*. New Jersey: Prentice-Hall. 1965. Pp. xii + 372. 54*s*.

WINNICOTT, D. W. *The Maturational Processes and the Facilitating Environment*. London: Hogarth Press. 1965. Pp. 295. 42*s*.

WOLMAN, B. B. & NAGEL, E. *Scientific Psychology*. London: Basic Books. 1965. Pp. xv + 620. 70*s*.

WOODSON, W. E. & CONOVER, D. W. *Human Engineering Guide for Equipment Designers*. London: University of California Press. 1965. 80*s*.

YINGER, J. *Towards a Field Theory of Behaviour*. Maidenhead: McGraw-Hill. 1965. Pp. ix + 365. 72*s*.

YOUNG, M. *Innovation and Research in Education*. London: Routledge and Kegan Paul. 1965. Pp. viii + 184. 25*s*.

YOUNG, R. K. & VELDMAN, D. J. *Introductory Statistics of the Behavioural Sciences*. New York: Holt, Rinehart and Winston. 1965. Pp. x + 345. $7.95.

ZALEZNIK, A. & HOMENT, D. *Casebook on Interpersonal Behaviour in Organisations*. London: Wiley. 1964. Pp. xvi + 585. 60*s*.

ZAX, M. & STRICKER, G. *The Study of Abnormal Behaviour: Selected Readings*. New York: Macmillan. 1964. Pp. x + 447. 38*s*.

ZEIGARNIK, B. V. *Pathology of Thinking*. New York: Consultants Bureau. 1965. Pp. xi + 211. $12.50.

ZINBERG, N. *Psychiatry and Medical Practice in a General Hospital*. New York: International Universities Press. 1965. Pp. xii + 364. 68*s*.

ZUBIN, J. *et al*. *An Experimental Approach to Projective Techniques*. London: Wiley. 1965. Pp. xix + 645. 105*s*.

THE BRITISH JOURNAL OF
PSYCHOLOGY

EDITED BY

ARTHUR SUMMERFIELD

ASSISTANT EDITORS

KEVIN CONNOLLY PETER MITTLER

VOLUME 56, 1965

CAMBRIDGE
AT THE UNIVERSITY PRESS
1965

PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W. 1
American Branch: 32 East 57th Street, New York, N.Y. 10022

Printed in Great Britain at the University Printing House, Cambridge
(Brooke Crutchley, University Printer)

CONTENTS OF VOLUME 56

PART 1. FEBRUARY, 1965

PARTS 2 AND 3. AUGUST 1965

PART 4. NOVEMBER 1965

INDEX OF AUTHORS

INDEX OF PUBLICATIONS REVIEWED

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Psychology (Monograph Supplements)
The British Journal of Medical Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology (about 512 crown quarto pages per volume), *The British Journal of Medical Psychology* (about 350 crown quarto pages per volume), and the *British Journal of Social and Clinical Psychology* (about 320 crown quarto pages per volume) appear in quarterly parts, the subscription price of each, payable in advance, being £4. 10*s*. per volume post-free ($16.50 in the U.S.A.). Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, N.W. 1. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 288 pages. The subscription rate, payable in advance, is £3. 10*s*. per volume post-free ($13 in the U.S.A.). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, London, E.C. 4; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number. The subscription price in the U.S.A. is $4.20. Subscriptions should be sent to the Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36 Essex Street, Strand, London, W.C. 2.

A series of *Monograph Supplements* is issued in connexion with *The British Journal of Psychology* but is not included in the subscription for that journal. Orders may be sent to any bookseller, or the Cambridge University Press, 200 Euston Road, London, N.W. 1, from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor A. Summerfield, Department of Psychology, Birkbeck College, Malet Street, London, W.C. 1;

for *The British Journal of Medical Psychology* should be sent to

Dr T. Freeman, Dundee Royal Liff Hospital, By Dundee

for *The British Journal of Social and Clinical Psychology* should be sent to

Mr Michael Argyle (Social Psychology Editor), Institute of Experimental Psychology, 1 South Parks Road, Oxford, *or* to

Dr R. R. Hetherington (Clinical Psychology Editor), 6 Abercromby Square, Liverpool 7;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Dr R. J. Audley, Department of Psychology, University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Mr L. B. Birch, The Institute of Education, The University, Sheffield 10.

for *The Monograph Supplements* should be sent to

Professor I. M. L. Hunter, Department of Psychology, The University, Keele, Staffs., whose decision regarding the publication of all Monograph Supplements will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1 September of the preceding year.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £6. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1 annually), *The British Journal of Medical Psychology* (£1 annually), *The British Journal of Social and Clinical Psychology* (£1 annually), *The British Journal of Mathematical and Statistical Psychology* (£1. 10*s*. annually), and *The British Journal of Educational Psychology* (£1 annually).

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, Tavistock House South, Tavistock Square, London, W.C. 1.

The British Journal of Psychology *Volume 56, Part 4, 1965*

CONTENTS

Printed in Great Britain at the University Printing House, Cambridge